University of Zurich[UZH]

PATTERNS AND DYNAMICS OF HUMAN BEHAVIOR IN OPTIMAL STOPPING PROBLEMS

Cumulative Thesis

presented to the Faculty of Arts and Social Sciences

of the University of Zurich for the degree of Doctor of Philosophy

by

CHRISTIANE MONIKA BAUMANN

Accepted in the spring semester 2021

on the recommendation of the doctoral committee:

Prof. Dr. Klaus Oberauer, main supervisor

Prof. Dr. Bettina von Helversen, co-supervisor

Zurich 2021

ABSTRACT

Important decisions in everyday life require a sequential search of options, often without the possibility to return to an earlier choice. The optimal stopping strategy selects the first option that is above a sequence dependent threshold. Human decision making commonly diverges from this strategy but the reasons for this discrepancy remain unknown. The goal of this thesis is to help understand and predict human response in situations that require sequential decisions by exploring links between computational models and psychological processes.

The first part of this thesis introduces a novel, data driven, linear threshold model (LTM) for human stopping decisions in sequential decision making. Three studies demonstrate that the proposed LTM outperforms existing models. Furthermore, the results reveal that the LTM predicts accurately search behavior in different environments and in real–world problems.

In the second part, the thesis presents investigations on how people adapt to outcome variance and time horizon within the framework of the LTM. Optimal stopping problems appear in a variety of contexts, including changing time limits or different environments, and humans have to adjust the parameters of their strategy. The results demonstrate that people adjust to both task variables. We find that decision thresholds are perfectly adapted to the variance of the sampling distribution, indicating that the value of an option is perceived relative to the options within the sequence. Moreover the results reveal that people adjust parameters consistently

across time horizons. In particular, longer time horizons lead to more selective initial decision thresholds and weaker adjustments thereof during the course of search.

Several studies have found that people stop their search too early and suggested risk aversion as a potential explanation. However, studies attempting to link individual risk preferences elicited in single decisions with search behavior in sequential decision tasks have failed to show any relationship. The third part of this thesis investigates the reasons of this non-convergence, by identifying the characteristic components of sequential search tasks which may affect choice behavior. The results reveal that both the sequential order of choices and the unequal frequency of accept and reject decisions lead to systematic biases in people's choices, giving rise to distortions between observed and stable risk preferences.

The thesis contributes to the theory of psychology as well as to applied decision making. Understanding the psychological processes that underlie sequential decisions will help quantify the conditions under which people may succeed or fail in such tasks. In turn, this will provide knowledge necessary for structuring decision problems in order to assist humans in making critical decisions that reflect their fundamental individual preferences.

ZUSAMMENFASSUNG

Viele wichtige Entscheidungen unseres Lebens erfordern eine sequentielle Suche, oft ohne die Möglichkeit zu einer früheren Option zurückzukehren. Die optimale Strategie zu solchen sogenannten *Optimal Stopping Problemen* wählt die erst Option welche über einem sequenzabhängignen Schwellwert liegt. Das menschliche Entscheidungsverhalten weicht von solch einer Strategie ab, wobei die Gründe dafür im Dunkeln liegen. Das Ziel dieser Arbeit ist es, menschliche Antworten in sequentiellen Entscheidungen zu verstehen und vorauszusagen, indem Verbindungen zwischen Computermodellen und psychologischen Prozessen untersucht werden.

Der erste Teil dieser Dissertation führt ein neues daten-getriebenes Entscheidungsmodell ein, das linear threshold model (LTM), welches das Antwortverhalten in sequentiellen Entscheidungen beschreibt. In drei Studien wird gezeigt, dass das LTM bisherige Modelle der Optimal Stopping Literatur übertrifft. Weiter wird nachgewiesen, dass das LTM menschliches Verhalten nicht nur in neuen Umgebungen, sondern auch in realistischen Suchproblemen beschreiben und voraussagen kann.

Im zweiten Teil dieser Arbeit wird untersucht, wie sich Menschen an die Streuung der Alternativen und an Zeithorizonte anpassen. Optimal Stopping Entscheidungen treten in unterschiedlichen Kontexten auf und die Parameter der Strategie müssen angepasst werden. Resultate zeigen, dass Menschen ihre Entscheidungen an beide Aufgabenvariablen angleichen. Die Entscheidungslevels sind perfekt an die Streuung der Optionen angepasst was darauf hindeutet, dass der Wert einer Option in

Optimal Stopping Aufgaben relativ zu den übrigen Alternativen bestimmt wird. Die Ergebnisse der zweiten Studie ergeben, dass Parameter konsistent an den Zeithorizont angepasst werden. Zudem führen längere Sequenzen an Alternativen zu einem höheren initialen Anspruchsniveau und einer schwächeren Anpassung über die Suche hinweg.

Viele Studien haben gezeigt, dass die Suche in Optimal Stopping Aufgaben zu früh beendet wird. Dieses Verhalten wird mit Risikoaversion in Verbindung gebracht. Allerdings haben neuere Studien ergeben, dass Risikopräferenzen gemessen in sequentiellen Entscheidungsaufgaben nicht mit Risikomassen aus einzelnen Risikoentscheidungen korrelieren. Im dritten Teil dieser Arbeit wird geprüft, inwieweit die spezifischen Komponenten der sequentiellen Aufgabe das Entscheidungsverhalten beeinflussen, welche zu dieser Nicht-Konvergenz führen. Die Ergebnisse deuten darauf hin, dass sowohl die sequentielle Präsentation wie auch das hohe Ungleichgewicht an Entscheidungen zu akzeptieren oder abzulehnen das Entscheidungsverhalten systematisch verändern und zu Abweichungen zwischen beobachteten und stabilen Risikopräferenzen führen.

Diese Dissertation leistet einen Beitrag zur Theorie der Psychologie wie auch zur angewandten Entscheidungsfindung. Das Verständnis der zugrundeliegenden Prozesse in sequentiellen Entscheidungen hilft zu quantifiziern, unter welchen Bedingungen die Menschen Erfolg haben oder scheitern. Dieses Wissen wiederum hilft sequentielle Entscheidungen so zu strukturieren, dass man bei diesen wichtigen Entscheidungen unterstützt wird.

ACKNOWLEDGEMENTS

I am grateful to Professor Bettina von Helversen for guiding me throughout this thesis. Bettina has been extremely generous with her knowledge and I very much appreciate the kindness and empathy that she offered me. I am indebted for her support and guidance and for all that she taught me. Bettina made this journey possible for me and I consider myself blessed and proud to have her as a mentor. Professor Klaus Oberauer served as an advisor during the last stage of this thesis with generosity and kindness. I was privileged to be part of his group at the University of Zurich. Professor Sam Gershman (Harvard University) hosted me during the first year of my thesis in his wonderful laboratory. It has been an invaluable experience that shaped my thinking.

Agnes Rosner lend me a friendly ear and shoulder, shared ups and downs. I thank her for deep and light conversation and for being the best possible travel companion. René Schlegelmilch took time to mentor me in new skills and was my ever present sounding board for ideas and questions. I thank him for many enlightening conversations and most of all his friendship. Henrik Singmann made the posterior predictive p-values a part of my everyday life, and I am thankful for our constructive and supportive collaboration. Andra offered wonderful insights on classical music as well as all other musical aspects of life. I thank my colleagues from the Cognitive Psychology Unit, Clara, Eda, Anke, Dave, Samuel, Gidon, Julia, Tuule, and Lea for the pleasant time. Nicole and Rahel offered invaluable help with administrative issues throughout my studies. Vassilios Kaxiras was a good wizard that dispelled many bugs, and Fabian Wermelinger made the thesis finally possible by helping me with its layout.

A very special hug to my sister Dominique for her selfless encouragement and care, my sister Kiki for taking care of my kids when I needed it the most, and for showering them with love and my mother for always being there for me and putting all decisions into perspective.

Dimitris and Monica were patient, accepting pasta as the optimal choice for dinner for the n*th* time in succession and offered unstoppable tender hugs that kept me going. Finally, to Petros: without you believing in me and standing by me, I would have never made it. You gave me support, discussed ideas and filled empty moments with warmth and love.

CONTENTS

INTRODUCTION

Some of the most important decisions in a human life, ranging from finding a partner or a job to choosing an apartment or a parking spot, require a sequential processing of decision options. Sequential decision making is necessary either because options are separated in time (Dudey et al., 2001) or they reside in a dynamic and uncertain environment (Paz et al., 2016). Decisions in such problems involve a tradeoff of accepting a subprime option prematurely and the danger of rejecting the best offer out of false hopes for the future.

For instance, imagine you plan an extended research visit in Boston and have to find a place to stay in a particular neighborhood. Apartments are rented out quickly and once you refuse an offer, it might not be available anymore. Your search is limited in time and thus you have to accept or reject a place without knowing if future apartments will be more attractive.

This class of optimization problems, generically known as optimal stopping problems (see Ferguson, 1989, for a historical overview), have features that make them well-suited to studying human decision-making on limited sequences of alternatives. For this reason, these problems have received significant and sustained attention by theoretical and empirical studies in cognitive psychology (e.g., Bearden et al., 2005; Corbin et al., 1975; Dudey et al., 2001; Kahan et al., 1967; Lee, 2006; Rapoport et al., 1970; Rydzewska et al., 2018; Seale et al., 1997, 2000; von Helversen et al., 2012; von Helversen et al., 2011) and other fields, such as experimental economics (e.g., Cox et al., 1989; Kogut, 1990; Zwick et al., 2003).

While significant research has focused on comparing human behavior to optimal solutions (Bearden et al., 2006; Gilbert et al., 1966; Seale et al., 1997), the psychological processes underlying them have been largely neglected. Furthermore, the focus on comparing behavior to an optimal solution has let researchers to constrain the structure of the decision problem in order to facilitate the computation of the optimal solutions. Some of these constraints, such as assuming that people have no knowledge about the distribution of the quality of the options have distanced decision problems in a laboratory setting from human behavior in the real world.

The "secretary problem" has been the most representative task for optimal stopping problems. In a typical scenario, the aim is to find the "best" of N applicants for a secretary job with optimality (the "best") measured by a single criterion/objective. The N applicants are presented in a sequence and the searcher has no knowledge of the distribution of the applicants' quality. The searcher must decide whether to hire the newly presented applicant or to continue the search in hope for a better applicant. She cannot choose an applicant that was earlier rejected (e.g., Bearden et al., 2006). The optimal solution for this problem involves searching for 37.5% of the available applicants and then taking the next one that is better than the ones seen so far (Ferguson, 1989; Gilbert et al., 1966). Indeed, behavioral research has generally found that people's behavior can be well described by such a cut-off strategy, however people tend to set their threshold too low (Bearden et al., 2006; Seale et al., 1997; von Helversen et al., 2011).

Some studies have investigated tasks closer to real sequential choice problems in which option values are drawn independently from a known distribution and the decision maker observes this exact values as each option is considered. (Guan et al., 2018; Guan et al., 2015; Kogut, 1990; Lee, 2006; von Helversen et al., 2012). In this version, the optimal solution is to choose according to a thresholds which is based on the probability of winning on the later positions (Gilbert et al., 1966, Section 3). Corresponding to the optimal approach, it has been argued that whether people accept or reject an option is linked to an internal threshold or aspiration level. If an option is better/worse than the threshold, it is accepted/rejected. However, there are differences in whether individuals have a single fixed threshold or the threshold decreases as the sequence progresses (Lee et al., 2014; von Helversen et al., 2012).

Several of these studies (Guan et al., 2018; Guan et al., 2015; Kogut, 1990; Lee, 2006) have the restriction that only the best alternative is rewarded—a payoff function that is found in only few real world situations. We often find a place to live in, a partner, or a secretary to hire, and thus receive a payoff, even if it is not the highest possible one. In this variant of the search problem, the optimal decision thresholds are calculated based on the expected reward of the remaining options (Gilbert et al., 1966, Section 5b) and change dynamically across position. However, research has shown that in the presence of such complex tasks, the innate bounded rationality (Simon, 1997) limits the ability to memorize and process all information. In turn this makes us rely on simplified mental models (Camerer, 1998; Gigerenzer et al., 2002). Simon (1955, 1957) has argued that rather than maximize, people often satisfice when making decisions. Satisficers have an internal threshold of acceptability against which they evaluate options, and will choose a decision outcome when it crosses this threshold. Therefore, satisficers are content to settle for a good enough option—not necessarily the very best outcome in all respects.

The goal of this dissertation is to help understand and predict how people respond in situations that require sequential decisions by developing a computational model and linking it with psychological processes. While focusing on realistic optimal stopping tasks, where payoff is proportional to the chosen value, I implement a model and conduct a comparative study of its predictions with those of human decisions. Furthermore, I will use this model to understand how people are affected by variations in the task context and by specific characteristics of the task itself.

In the first part of thesis, I introduce a threshold model of human decision behavior in optimal stopping tasks that is an instantiation of the satisficing heuristic proposed by Simon (Simon, 1955). The *linear threshold model* (LTM, Baumann et al., 2020) proposes that humans rely on a mental shortcut and use internal decision thresholds that are adjusted in a linear manner across the sequence of search. The model identifies and measures three hidden processes underlying human choices in an optimal stopping task: an initial aspiration level, its adjustment across the sequence and the choice sensitivity towards deviations between the option and the (adapted) aspiration level. Results reveal that the LTM is capable to accurately predict human choices in optimal stopping paradigms and further generalizes to changing environments.

The desire to determine cognitive mechanisms in optimal search task has led me to the second research question of this thesis: How do people adapt to changes in the task structure and how is this adaptation reflected in the cognitive parameters? The context of the decision problem (number of alternatives, time pressure) demands from the decision maker to adjust her search strategy in order to succeed. Despite the consensus that people adjust their choices to structural differences in optimal search tasks (Corbin et al., 1975; Cox et al., 1989; Guan et al., 2018; Lee et al., 2004; Shapira et al., 1981), the question of *how* people adjust is mostly neglected. The goal of the second part is to address the reasons and the robustness of adaptive behavior in optimal stopping problems. The analysis relies on the LTM which provides a framework to measure and identify the processes involved in human adaptation.

Despite the results showing stable individual differences in cognitive parameters across structurally changing tasks, people are very heterogeneous with respect to their choice behavior. It has been suggested that this heterogeneity in dynamic choice situations is reflected in risk preference heterogeneity (Schotter et al., 1981; Schunk, 2009; Sonnemans, 1998, 2000) whereas risk averse agents would stop earlier in the sequence. However, studies attempting to link individual risk preferences elicited in single gamble tasks with search behavior in sequential decision tasks have failed to show any relationship (Frey et al., 2017; Pedroni et al., 2017; Schunk, 2009; Schunk et al., 2009). Moreover, the studies throughout this thesis have revealed a time dependence for the participants risk attitudes. While at the beginning of the sequence, participants accept a choice too often compared to the optimal model, thus seem to behave risk averse, they become too selective at the end of search, which corresponds to apparently risk seeking behavior. The core of the asymmetry in risk preferences across the sequence resides in the linear structure of the decision threshold.

The third project in this thesis aims to determine the characteristic components if an optimal stopping tasks such as the stepwise incremental decisions or the fact that only one option can be chosen and measures their impact on the emergence of the linear threshold strategy. The understanding of systematic effects on the particular task characteristic on human search behavior may reconcile the seemingly inconsistent findings in risk preferences.

1.1 MODELS OF OPTIMAL STOPPING BEHAVIOR

The optimal and the linear threshold model are introduced based on a classical optimal stopping example which is employed as experimental paradigm throughout the thesis. We consider a decision maker (here a customer) who is planing a vacation and chooses to buy the plane ticket online. The ticket prices vary from day to day and the customer wants to find the cheapest price. The customer checks the offer every day and decides if she wants to accept or reject it, without having the option to go back in time to a previously rejected price. Search time is limited by her vacation schedule (i.e., 10 decisions per trial) and, once accepted, the search ends.

More formally, we consider a decision maker who encounters a sequence of prices with values denoted by $x_1, \quad ,x_{10}$ and she wants to find the minimum value in the sequence. If the decision maker accepts the price x_i, the sequence terminates and she has to pay x_i; otherwise, she continues to the next ticket. When the last ticket is reached, it must be accepted.

1.1.1 *Optimal stopping model*

The optimum strategy is derived based on the expected reward of the remaining prices, which depends upon the distribution of the prices' values as well as upon the number of price opportunities. Based on the expected reward, a threshold T_i is calculated for each day i (for a detailed description of the calculation of optimal thresholds, I refer to the Supporting Material of Chapter 2: A linear threshold model for optimal stopping behavior, Text A). Consequently, a price x_i is accepted

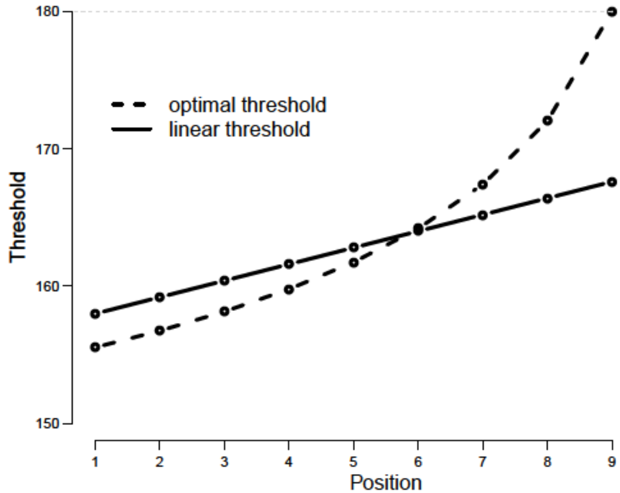

FIGURE 1.1: Optimal and linear thresholds when options are sampled from $\sim \mathcal{N}(\mu = 180, \sigma = 20)$ and the goal is to find the minimum. Optimal thresholds: dotted line, linear thresholds: solid line, mean of the sampling distribution:light dotted line

if it is below the position dependent threshold T_i^* and rejected otherwise. Figure 1.1 displays the non-linearly increasing optimal thresholds for a sequence of 10 alternatives.

1.1.2 *Part 1: Linear Threshold Model*

The optimal solution to the problem is very complex and requires to use backward induction, which has been shown to be very challenging even for a few time steps (Huys et al., 2015). Moreover, research has found that human choices are not fully described by assuming that they use such optimal thresholds when deciding to accept or reject an alternative (e.g. Goldstein et al., 2020; Guan et al., 2014; Lee, 2006;

von Helversen et al., 2012) and suggest that people must rely on a different strategy when solving such tasks. I propose the *linear threshold model* (LTM, Baumann et al., 2020) which postulates that human's decision thresholds are constrained by a linear relation to each other. In particular, it assumes that human choices in optimal stopping problems can be described by (1) an aspiration or acceptance level which reflects the option's value above which an option is accepted directly in the first decision (2) an adaptation rate, reflecting the change of the aspiration level with the time/number of available options running out in a linear manner and (3) choice sensitivity (or determinism), reflecting how sensitively people react to the extent by which the current option value deviates from the (adapted) aspiration level.

More formally, the linear decision thresholds are completely defined by the first threshold T_0 and the linear increase as the sequence progresses,

$$T_{i \quad 1} \qquad T_i \qquad i. \tag{1.1}$$

The model further assumes that the decision maker relies on a probabilistic threshold to make the decision to accept or reject a ticket–i.e., ticket x_i on position i is compared to a position dependent threshold T_i. This comparison yields an acceptance probability $_i$ based on a sigmoid choice function with sensitivity parameter and

$$_i \qquad \frac{1}{1 \quad \exp \qquad x_i \quad T_i}. \tag{1.2}$$

This model entails three free parameters, the first threshold T_0, the increase of the threshold , and the choice sensitivity .

Comparing the LTM to existing models of sequential decision making showed that it provided a better account for the observed data. Furthermore, it provided an overall excellent account of the data. More specifically, we calculated – on a by-participant basis – posterior predictive p-values comparing the misfit between the observed data and the model predictions with the misfit between synthetic data generated from the models and the model predictions. This analysis showed that for almost all participant the observed misfit was in line with the misfit expected under the model, indicating that the model adequately described their responses.

The LTM's ability to generalizes to new environments is crucial to be a useful model. Therefore, simulation studies from the LTM and the optimal model were conducted to predict choice behavior in scarce (values sampled from left skewed distribution) and plentiful environments (values sampled from right skewed distribution). While these models make contradictory predictions on search length, participants search behavior agreed with the LTM's prognosis. Whereas in the scarce environment, in which good values occur very rarely, participants searched too long compared to an optimal agent, they stopped too early in the plentiful environments. The third experiment extended the experimental paradigm to a real-life scenario, in which participants were asked to search sequentially for the best price for well–known commodity products collected on Amazon.com. The prices for the products varied in their mean and variance and thus thresholds were compared on a normalized scale. The results revealed a close agreement between participants' choice behavior and the LTM's prediction, indicating that the linear threshold strategy is a robust way to adjust to optimal stopping tasks.

1.2 ADAPTIVE BEHAVIOR IN OPTIMAL STOPPING TASKS

There is a broad agreement in the decision making literature that the processes used to solve decision problems vary as a function of a number of task and context factors (Payne, 1982; Payne et al., 1993; E. U. Weber et al., 2009, for reviews). Research on optimal stopping behavior has consistently shown that people are responsive to the environmental distributions of the values (e.g. Corbin et al., 1975; Cox et al., 1989; Guan et al., 2018; Kahan et al., 1967; Lee et al., 2004; Rydzewska et al., 2018; Shapira et al., 1981) or amount of alternatives within the sequence (Guan et al., 2020; Lee et al., 2004; Seale et al., 1997). However, despite the consensus that people adjust their choices to different contexts in optimal stopping tasks, relatively little is known about how adaptation occurs. Therefore, in the second and third part of this thesis, I address the issue of the reasons and robustness behind the adaptive behavior in optimal stopping problems.

More specifically, the second project investigates adaptive behavior to variations of task factors, such as the number of alternatives available and their dispersion. The analysis relies on the LTM which provides a framework to study the differential effect of adaptation in respect to the underlying cognitive variables. The third project examines characteristic components of an optimal stopping task that give rise to people's linear threshold strategy which in turn leads to apparent inconsistencies in risk preferences.

1.2.1 *Part 2: Search environment*

1.2.1.1 *Variance*

Does the variation of alternative prices affect the choice to accept or reject a particular offer? Research in simultaneous decision making has repeatedly shown that valuation of an option is sensitive to the range of other options (e.g. Nieuwenhuis et al., 2005; Padoa-Schioppa, 2009; Rigoli et al., 2016; Tversky et al., 1993). Moreover, some researchers suggest that people value an alternative based on the relative rank within a pool of alternatives, and not by their absolute value (e.g. "normalization hypothesis", Rangel et al., 2012). Under this hypothesis, people are expected to transform their value representations into value rankings taking the variance of the alternatives into account. Decision thresholds are thus expected to differ on an absolute but not on a normalized scale.

Alternatively, research on risky decision making suggests that variance in the outcomes changes people's risk preferences, where higher outcome variance implies higher risk and thus results in more risk averse behavior (Genest et al., 2016; Holt et al., 2002; Markowitz, 1959; E. Weber et al., 2004). As a consequence, acceptance rates would be increased in high variance environments, resulting in shorter search length and lower performance. Adaptive behavior to outcome variance would thus be reflected in higher decision thresholds and potentially also higher adaptation rate, when expressed on a normalized scale.

The first study demonstrates that the decision thresholds in a low and a high variance environment are practically identical on a normalized scale (0, 1), indicating that people accept prices within the same percentile on each position.

This result suggests that the price's value is assessed on the basis of its percentile rank within the sample rather than on the basis of its absolute number. This finding is in line with the normalization hypothesis, which assumes that the value of an option is computed under a normalized code and corresponds to its relative position in the distribution of options (Rangel et al., 2012; Stewart et al., 2006).

1.2.1.2 *Time Horizon*

In optimal stopping task, time is a significant variable and must be included into the decision process. For example, having plenty of time to search for a good ticket price allows to increase the aspiration level whereas little time requires to lower your standards. Previous studies have reported that humans search more, in absolute terms, in longer time horizons (Lee et al., 2014; Seale et al., 2000) which is reflected in either a higher initial aspiration level or a weaker adjustment rate during search. To investigate the details of this phenomenon I used sequences with 5, 10 and 20 available choices (sequence positions). Thereby, three hypotheses were tested: If time horizon affects (a) the initial threshold, (b) the change of threshold over time, or (c) both.

Results revealed that the effect of time horizon was captured by changes in initial decision thresholds (T_0), and the adaptation rate ( ) of the thresholds with ongoing sequence positions (i.e., with growing number of rejections). More specifically, with longer time horizons, the initial aspiration level increased reflecting initially lower acceptance rates in longer time horizons. Simultaneously, the adjustment of the aspiration level during search was reduced, so people were less willing to depart away from their initial decision threshold. Additionally, subjects adapted in a way that a LTM simulation indicated were appropriate given the changes in the structure of the decision task. Results further suggest that participant's initial thresholds and adjustment rates are adapted systematically to time horizons in a non-linear way, indicating that they may follow scaling regularities across time scales. Indeed, the best performing linear thresholds predict a proportional adaptation of the initial threshold and its adjustment rate and in this sense, participants' adaptation is optimal.

1.2.2 *Part 3: Characteristic components of sequential search task*

The linear adaptation of the decision thresholds across search has been shown to be a robust and efficient way to solve sequential search tasks. However, compared to the optimal thresholds, it reveals a consistent pattern: While thresholds are usually higher at the beginning of the search, thus people accept too often, they tend to be too low at the end, thus people accept too little. This behavior is inconsistent under the perspective of risk preferences. While the optimal model assumes expected value maximization, i.e., risk neutrality, participants choose as if they reverse their risk preference during the course of search, from risk averse to risk seeking behavior. In my third project I investigate if the characteristic structure of an optimal stopping task encourages the emergence of linearity which leads to the appearance of asymmetries in risky behavior. In particular, I tested in three experiments if (1) the presentation of the underlying sampling distribution (2) the stepwise incremental decisions or (3) the unequal frequency of accept and reject decisions leads to systematic biases in peoples' decision thresholds which give rise to the observed risk attitudes during search.

1.2.2.1 *Experience versus description*

In an optimal stopping task, subjects discover the sampling distribution by observing options from the respective distribution. However, several studies have shown that the way how we acquire information about a distribution can lead to a systematic bias in estimates thereof (Barron et al., 2003; Hertwig et al., 2004; E. Weber et al., 2004). In particular, experimental investigation on decision under risk relies on two distinct types of paradigms, involving either experience- or description-based choices (Hertwig et al., 2009). In experience-based choices, outcome distributions are inferred by encountering repeated draws from the respective distribution, which corresponds to the optimal stopping paradigm. In decisions by description, outcome distributions are explicitly described, by numerically presenting outcomes and probabilities. Furthermore, it was suggested that outcome and probability information translate into systematically different subjective representations in description- versus experience-based choice (Hertwig et al., 2004; Lejarraga et al., 2011; Madan et al., 2019; Wulff et al., 2018). This perspective leads to the assumption that biases in optimal stopping behavior may emerge from the sampling distribution's format. The first study thus compares choice behavior between two tasks which only varied in the presentation of the sampling distribution. It was hypothesized that the format of the presentation, by experience of by description, causes different biases in the estimates of the underlying distributions which would result in diverging search behavior.

Results showed no significant difference in performance between the two conditions, despite the fact that participants accepted earlier than in the descriptive format. Importantly, the way how the underlying distribution is presented – either by experience or by description – led to the same linear dependency between decision thresholds and thus to an apparent shift in risk preferences in both conditions. In contrast to the hypothesis, these results demonstrate that the presentation format of the underlying distribution has a minor effect on peoples strategy selection in optimal stopping task.

1.2.2.2 *Sequential versus single gambles*

People often have illusory beliefs about the sequential dependence of successive independent events. In the gambler's fallacy, negative dependence is assumed: A run of heads leads me to believe that tails is due. In the "hot hand" fallacy, a positive dependence is inferred: A sequence of successful shots leads me to believe that a basketball player is temporarily "in for" (Gilovich et al., 1985; Keren et al., 1994). In both cases, present payoffs are believed to differ from past payoffs, because event probabilities are assumed to have changed, although this is not in fact the case in these particular instances (Miler et al., 2009). Following this idea, decisions in optimal stopping tasks may be prone to the gambler's fallacy by believing that a sequence of low values must be followed by a high one. As a consequence, decision thresholds would depend on the prior history and not – as prescribed by the optimal rule – exclusively on the remaining future options. Consequently, this bias in decisions would – at least partly – overshadow stable risk preferences, leading to the apparent shift in their risk attitudes. The second study thus was designed to test if the sequentiality in choices has an effect on peoples decision thresholds. Participants performed two versions of the optimal stopping task which differed only in the presentation format: The first one corresponded to a regular optimal search problem while the second one consisted of single gambles which were identical to the first one in respect to outcomes and probabilities.

Results revealed that people's choices differed significantly between sequential and single choices, despite their identical underlying statistical properties. While the sequential task encourages the linear strategy, single choices facilitate decision thresholds to be updated more independently. Consequently, in the later case, the adaptation across the sequences is in closer agreement with the non-linear pattern predicted by the optimal model, which leads to an attenuation of the apparent risk asymmetries in optimal stopping tasks.

1.2.2.3 *Unequal frequency in accept and reject decisions*

A third important characteristic in optimal stopping tasks is that only one option can be accept within the sequence. Consequently, this leads to an imbalance between choices such that reject decisions are predominant during the session. While the optimal model prescribes to reduce the proportion of reject decisions during search, reaching 50% on the second–to–last position, people's reject rate is significantly below this rate at the end of search. Studies have found that people are inclined to repeat decisions which were favourable prior in history (Alós-Ferrer et al., 2016; Erev, 2012) and argue that there exists a strong tendency to simply reiterate the most recent decision, which is even stronger than the tendency to react optimally to the most recent outcome. From this viewpoint, peoples' higher reject rate on the later positions – compared to the optimal model – could arise from the overall dominance of reject decisions within the process of search. Despite this decision bias being unrelated to risk preferences, it may contribute to the seemingly increase of risk seeking behavior observed on final positions. To examine this possibility, participants performed two tasks: In the first one, the proportion of accept and reject decisions corresponded to the ratio found in optimal stopping tasks. In the second one, inequalities in accept and reject choices were removed.

Results show that the imbalance in choices affects choice behavior significantly. Once people decide in a setting in which the inequality is removed, decision thresholds are adjusted non-linearly across positions and in close agreement with the optimal model. This result indicates that an unequal proportion of opposite choices in the decision environment leads to biases in people's choices which at the same time overwrite stable risk preferences.

1.2.3 *General Discussion*

In many real life decisions, options are distributed in space and time, making it necessary to search sequentially through them, often without a chance to return to a rejected option. The psychological processes underlying sequential decisions are still poorly understood even though some of the most important decisions such as the choice of a job, or a partner are made sequentially. The goal of this thesis is to contribute to this topic by increasing knowledge about the psychological processes underlying sequential decisions and their responsiveness to changes in the context. In the first project, I introduce the linear threshold model, suggesting that humans decide on the basis of an internal decision threshold which is adjusted in a linear matter across the sequence. As a consequence, people become less selective during search, which may lead them to accept options on later time steps which were rejected earlier in search. Results revealed that participants' choices are well captured by the model and accurately predicted in new environments. Based on these findings, the second project addresses the question of robustness and adaptability of the cognitive parameters to task contexts, such as the number of alternatives or their dispersion. The manipulation of the sampling variance revealed that people adapt their decision thresholds on an absolute scale, while a higher variance leads to lower decision thresholds (when searching for the minimum). However, the difference in decision thresholds disappears on a standardized scale (0, 1), which leads to the conclusion that decision thresholds are formulated on a percentile level (e.g. accept all prices on the first position that belong to the 12% lowest prices within the sampling distribution). The manipulation of the number of alternatives resulted in adjusted choices in a way which the LTM simulation indicated were appropriate given the changes in the structure of the decision task. In particular, an increased amount of alternatives resulted in more selective initial decision thresholds and weaker adjustments across the sequence. Moreover, the study of the robustness of cognitive parameters provides evidence that the LTM parameters reflect stable individual processes across changing contexts.

Participants' linear decision thresholds revealed a consistent phenomenon throughout this thesis' studies. While acceptance rates are too high, compared to the optimal model, in the beginning of search, they are too low at later stages. This behavior reflects a shift in risk attitudes across search, from risk averse to risk seeking behavior. The third project examined if the particular characteristic of sequential search task, such as sequential presentation or the inequality of accept and reject decisions, could account for the linear thresholds and thus for the apparent asymmetries in risk preferences. The first experiment indicates that the presentation of the sampling distribution, by experience or by description, has no effect on the linear updating of decision thresholds, and thus leads to similar inconsistencies of risk preferences across search. However, the sequential order of choices and the unequal frequency of accept and reject decisions reveal a significant effect on decision thresholds, encouraging the emergence of linear threshold. Once these components are removed, decision thresholds are adapted independently and in closer agreement with the optimal model. This result indicates that the characteristics of a sequential search tasks leads to systematic biases in choice behavior which override stable risk preferences.

This thesis' results provide evidence that humans tend to adjust their decision thresholds in a linear matter to solve the complex task of optimal stopping search. Simulations studies have shown that the linear threshold strategy achieves an almost optimal performance and thus it represents an efficient way to adapt to the task. However, when sequences grow larger (i.e., 20 alternatives), participants struggle to adapt to the best performing linear thresholds and thus their performance decreases. Computational modelling results indicate that an increase in complexity leads them to diverge from a strict linear threshold strategy. Indeed, research has suggested that humans' strategy selection can be formulated as a function of both its costs, primarily the effort required to use a rule, and its benefits, primarily the ability of a strategy to select the best alternative (Beach et al., 1978; Payne et al., 1988; Russo et al., 1983). Under this perspective, peoples' divergence from a strict linear threshold rule in extended time horizons may be caused by an increased effort outweighing its benefits.

The value of an option in an optimal stopping task seems to be encoded as its percentile rank within its sampling distribution. Consequently, decision thresholds are formulated on percentile levels (e.g, accept all prices that belong to the 12% lowest prices within the sampling distribution). This finding is in line with studies showing that people evaluate an attribute upon their relative rather than their absolute value within the generating distribution (Rangel et al., 2012; Stewart et al., 2006). For example, Brown et al. (2008) argue that ordinal rank has a statistically significant effect upon well-being and that to understand what make human beings content is is therefore necessary to look at the whole distribution of incomes. This thesis' results have shown that despite the heterogeneity in individual decision thresholds, they are relatively consistent across task contexts. It would be interesting to investigate whether this observed individual differences in acceptance levels are related to more traditional measures of problem solving ability and psychometric intelligence. Given that optimal stopping problems are representative of a class of real world sequential decision-making tasks, they allow the possibility that there is a similar relationship for non-perceptual tasks to be examined.

This thesis has revealed the characteristic components of an optimal stopping problem involves that decision thresholds are updated depending on earlier outcomes. In particular, it seems as if the decision maker thinks that a good offer must follow a sequence of unsatisfying options. This a mental representation of the optimal stopping environment would correspond to drawing from a pool of options without replacement (i.e., hypergeometric distribution). Indeed, studies on related paradigms, such as the Balloon Analog Risk Task (BART, Lejuez et al., 2002) or the Iowa Gambling Task (Bechara et al., 1999), showed that people misunderstood the dynamic event probabilities in related paradigms (Pleskac, 2008; Wallsten et al., 2005). Preliminary results from a simulation study indicate that a distorted mental representation of the dynamic distribution may strongly contribute to the emergence of linearity in peoples' decision thresholds.

An optimal stopping choice represents a risky decision between a certain outcome and the risk to continue search for a better one. However, studies attempting to link individual risk preferences elicited in single gamble tasks with search behavior in sequential decision tasks have failed to show any relationship (Frey et al., 2017; Pedroni et al., 2017; Schunk, 2009; Schunk et al., 2009). Nevertheless, this work has found a relationship between risk preferences measured in the gamble task and search length in the optimal stopping task such that participants who were relatively more risk seeking in the single gambles searched more in the optimal stopping task. It seems that the specific implementation of the optimal stopping task used in this thesis allows to find convergence in risk preferences between single gambles and dynamic search tasks and thus may help uncover a common mechanisms underlying these risky decisions (Frey et al., 2017).

Limitations

While one of the strengths of this work is that it proposes a framework for understanding human decision making in optimal stopping problems, it will be imperative to investigate the limits of this framework in additional contexts. The work presented here is merely a first step toward answering important questions about underlying processes, and all three parts make way for a number of new questions to be answered. The experiments we present are simple and controlled in order to act as proof of concepts, but this factor limits the number of questions we can answer and does not always provide a perfect analogy to real-world circumstances.

1.2.4 *Future directions*

The LTM proposes that decisions in optimal stopping tasks are taken based on decision thresholds which are defined by the initial aspiration level and its linear adjustment on each position. Despite the LTM's accuracy in capturing people's choices, it assumes no change on the part of the participant as he or she proceeds through the session. This assumption is reasonable as long as the sampling distribution is known – either by providing an accurate description beforehand or by expecting participants to hold an accurate representation of the real-life distribution in long-term memory. However, in unknown environments participants have to gain an understanding of the environment during search and adapt decision thresholds accordingly. In such novel scenarios, modeling how the participant learns from experience is an important part in understanding the optimal stopping behavior. Bayesian learning models have been proposed in related sequential search tasks such as the BART (Lejuez et al., 2002), and the Iowa Gambling Task (Bechara et al., 1999). These models suggest that the decision maker uses Bayesian updating to change prior estimate of the payoff probabilities into posterior probabilities on the basis of experience with the outcome observed (Speekenbrink et al., 2015; Steyvers et al., 2009). Such algorithms are very promising candidates for incorporating learning and adaptation processes into the framework of the LTM.

The optimal stopping paradigm involves that one can not return to a previously rejected option which is in conflict with the properties of many real-world tasks. Many search situations involve the probability of going back to a former option or considering some of the options simultaneously. For instance, when deciding between job offers one may wait to make a decision about one offer until having seen the next one. However, this may imply that the probability of receiving this offer will be lower than when it was first received. Future work should extend the optimal stopping problem by allowing the decision maker to return to previously passed options and test how this setting would affect the linearity in decision thresholds. Furthermore, this analysis could reveal if choices improve – because now an erroneously passed option can be retrieved – or whether it may impair choices – because the higher complexity makes the task more difficult and takes more time.

Optimal stopping tasks involve stepwise incremental decisions and immediate feedback, which correspond to "hot" decision making paradigms. Figner et al. (2009) have shown that in a "hot" version of a dynamic risk taking decision task, the affective system tends to override the deliberative system in states of heightened emotional arousal. Therefore, affective processes may play an important role during the course of search, in which each time step involves an increase of risk. Indeed, this thesis has shown that the removal of the stepwise decisions leads to a change in peoples' decision strategy closer to the optimal solution, indicating more deliberative information processing takes control. It seems that the interplay of deliberative and affective processes plays a key role in sequential search tasks which should be addressed in future studies.

1.2.5 *Conclusion*

This thesis introduces a model of human behavior in optimal stopping tasks which proposes that decision thresholds are adjusted linearly over time. The model provides an accurate fit and makes non-trivial predictions about search behavior, which were confirmed experimentally. Furthermore, it investigates whether and how context task variables affect the linear model's cognitive parameters, and test their reliability. Results show that people consistently adapt their decision thresholds to variance and time horizon and indicate that thresholds are formed based on percentile ranks rather than absolute values. Finally, the thesis shows that peoples' linear threshold strategy is an adaptation to the characteristic components of an optimal stopping task and overrides to large extent their basic risk preferences. While there is an enormous experimental literature on the foundations of decision behavior in static decision situations, the foundations of behavior in dynamic decision situations, despite being equally important, remain largely unexplored. This thesis aims to shed light on humans' strategy, adaptivity and reliability in optimal stopping tasks and thus provides an important step in understanding human behavior in dynamic choice situations.

A LINEAR THRESHOLD MODEL FOR OPTIMAL STOPPING BEHAVIOR

Authors Christiane Baumann, Henrik Singmann, Samuel J. Gershman and Bettina von Helversen

Journal Proceedings of the National Academy of Sciences of the United States of America (2020)

DOI https://doi.org/10.1073/pnas.2002312117

ABSTRACT

In many real life decisions, options are distributed in space and time, making it necessary to search sequentially through them, often without a chance to return to a rejected option. The optimal strategy in these tasks is to choose the first option that is above a threshold that depends on the current position in the sequence. The implicit decision making strategies by humans vary but largely diverge from this optimal strategy. The reasons for this divergence remain unknown. We present a new model of human stopping decisions in sequential decision making tasks based on a linear threshold heuristic. The first two studies demonstrate that the linear threshold model accounts better for sequential decision making than existing models. Moreover, we show that the model accurately predicts participants' search behavior

in different environments. In the third study, we confirm that the model generalizes to a real-world problem, thus providing an important step towards understanding human sequential decision making.

2.1 INTRODUCTION

Decisions that arise in everyday life often have to be made when options are presented sequentially. For example, searching for a parking spot, deciding when to take a vacation day, or finding a partner, all require that the decision maker accepts or rejects an option without knowing if future options will be more attractive. Decisions in such problems involve a trade-off between accepting a possibly suboptimal option prematurely and rejecting the current offer out of false hopes for better options in the future.

Despite the importance of such decisions, relatively little work has been made toward characterizing the process by which humans decide to stop searching in natural settings of this task.

Earlier research has focused on a simplified version of optimal stopping problems, the so-called secretary problem, where only the rank of the option relative to those already seen is shown (Bearden et al., 2006; Seale et al., 1997, 2000) and only the overall best alternative is rewarded. In the secretary problem, the optimal strategy is to ascertain the maximum of the first 37% options and choose the next option that exceeds this threshold (Gilbert et al., 1966). Empirical studies suggest that in general people follow a similar strategy but usually set the cut-off (i.e., from which point on they will accept an option that exceeds the previous options) earlier than the 37% prescribed by the optimal solution (Kahan et al., 1967; Seale et al., 1997).

Some studies have investigated tasks closer to real sequential choice problems by presenting the actual value of the option to the decision makers (Guan et al., 2018; Guan et al., 2015; Kogut, 1990; Lee, 2006; von Helversen et al., 2011). In this version, the optimal is based on calculating the probability of winning on the later positions. From this probability, a threshold is calculated for each option in the sequence as described by Gilbert and Mosteller (Gilbert et al., 1966, Section 3). Lee (2006) estimated a family of threshold-based models and showed that most participants decreased their choice thresholds as sequences progress. Although people are overall quite heterogeneous in their search behavior, they tend to cluster around the optimal solution (Guan et al., 2018; Guan et al., 2015). Importantly, these studies still kept the restriction that only the best alternative is rewarded—a payoff function that does not correspond well with everyday experiences. Humans do find a mate, an apartment to live, or a ticket to fly to their vacation destination, and thus receive some payoff, even if that may not be the highest possible payoff.

In the present research, we propose a model of human decision making in optimal stopping problems using payoffs that are based on the actual values. In this variant of the search problem, the optimal decision thresholds are calculated based on the expected reward of the remaining options ((Gilbert et al., 1966, Section 5b) and Supporting Material, Text A). This leads to a decision threshold that changes notably nonlinear over the sequence.

In contrast, we propose that people rely on a mental shortcut and adapt their thresholds linearly over the sequence. We show that a model with this linearity assumption accurately captures when people stop search and accept an option, even in a real-world setting. Furthermore, this model allows us to predict under which conditions people search more or less than the optimal model, making it a useful tool to understand human sequential decision making.

We first sketch a family of cognitive models for describing behavior in optimal stopping problems. We then present results from three behavioral experiments that provide evidence for the validity of the linear model in a laboratory setting as well as in a real-world scenario.

2.1.1 *Computational models*

We explain the computational models based on a typical optimal stopping problem that we also used in our first two experiments. The decision maker (here a customer) is planing a vacation and decides to buy the plane ticket online. Ticket prices vary randomly from day to day and the customer wants to find the cheapest ticket. The customer checks the ticket price every day and decides if she wants to accept or reject the ticket, without having the option to go back in time to a previously rejected offer. Search time is limited by her vacation schedule (i.e., 10 decisions per trial) and, once accepted, the search ends.

More formally, we consider a decision maker who encounters a sequence of tickets with values denoted by $x_1, \quad , x_{10}$ and she wants to find the minimum value in the sequence. If the decision maker accepts ticket x_i, the sequence terminates and she has to pay x_i; otherwise, she continues to the next ticket. When the last ticket is reached, it must be accepted.

All models assume that the decision maker relies on a probabilistic threshold to make the decision to accept or reject a ticket—i.e., ticket x_i on position i is compared to a position dependent threshold t_i. This comparison yields an acceptance probability $_i$ based on a sigmoid choice function with sensitivity parameter and

$$_i \quad \frac{1}{1 \quad \exp \quad x_i \quad t_i}. \tag{2.1}$$

Small values of produce more stochasticity in decisions, whereas the policy approaches determinism when .

We examine the setting of thresholds by comparing the performance of four different models.

- The *Independent Threshold Model (ITM)* serves as our baseline. It assumes no dependency between the thresholds. It entails N independent threshold parameters $t_1, ..., t_N$, one for each position in the sequence, where the decision maker can decide to accept or reject an offer (at position $N \quad 1$ the ticket must be accepted). The thresholds can take any value across positions. The model maintains maximal flexibility and provides an upper limit how well any threshold model can describe a person's decision given the assumption of a probabilistic threshold.
- The *Linear Threshold Model (LTM)* postulates that the thresholds are constrained by a linear relation to each other and therefore are completely defined by the first threshold t_0 and the linear increase as the sequence progresses:

$$t_{i \quad 1} \quad t_i \qquad i, \tag{2.2}$$

This model entails three free parameters, the first threshold t_0, the increase of the threshold and the choice sensitivity .

- The *The Biased Optimal Model (BOM)* is based on the Bias-from-Optimal threshold model proposed by Guan et al. (Guan et al., 2015), assuming that humans are using thresholds that deviate systematically from the optimal thresholds.. The optimal thresholds t_i for each position i are derived by determining the expected reward of the remaining options (derivation in (Gilbert et al., 1966, Section 5b) and in Supporting Material, Text A). The model entails a systematic bias parameter that reflects the divergence of the human threshold from the optimal one. Additionally, the thresholds depend on a parameter that determines how much their bias increases or decreases as the sequence progresses.

$$t_i \quad t_i \qquad\qquad i, \tag{2.3}$$

When and are set to 0, the thresholds represent the optimal thresholds that lead to best performance. This model is therefore defined by three free parameters, , and the choice sensitivity .

- The *Cut-off Model (CoM)* is inspired by the optimal decision rule for the rank information version of the secretary problem where the distribution of the prices is unknown. It assumes that the DM has a fixed cut-off value k that determines how long she explores in the beginning of the sequence. The highest value seen in that initial sample of k tickets is then set as her threshold, and the first value that exceeds this threshold in the remainder of the sequence is chosen. This model has two free parameters, the cut-off value k and the sensitivity parameter .

Models were implemented in a *hierarchical-Bayesian statistical framework* using JAGS (Plummer et al., 2003) (Supporting Material, Text B).

2.2 EXPERIMENT 1: MODEL COMPARISON

We asked 129 participants to solve a computer-based optimal stopping problem following the ticket-shopping task described above. Tickets were normally distributed with a mean value of \$180 and a standard deviation of \$20. In the first phase, subjects learned the distribution using a graphical method proposed by (Goldstein et al., 2014) (*Material and Methods*). Supporting Material, Fig. 2.4 shows that this procedure was successful in ensuring participants learned the distribution.

In the second phase, participants performed 200 trials of the ticket-shopping task. In each trial, participants searched through a sequence of ten ticket prices. For each ticket, they could decide to accept or reject it at their own pace. Participants were aware that they could see up to 10 tickets in each trial, and they were always informed about the actual position and the number of remaining tickets (Supporting Material, Fig. 2.5 for a screen shot). It was not possible to go back to an earlier option after it was initially declined. If they reached the last ticket ($10th$) they were forced to choose this ticket. When participants accepted the ticket, they received feedback about how much they could have saved if they had chosen the best ticket in the sequence. Performance was incentivized based on the value of the chosen ticket (*Material and Methods*).

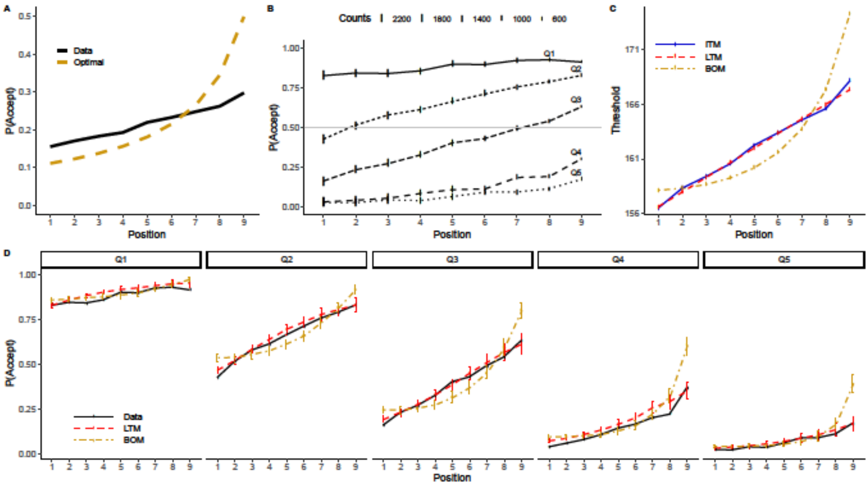

FIGURE 2.1: (A) Probability to accept a ticket on each position across all prices. The dark line represents participant's frequency to accept, the dashed yellow line an optimal agent's probability to accept. (B) Participants' probability to accept. Each line represents ticket prices ranging from the first quantile to the fifth quantile. Q1: Tickets in first quantile, Q2: Tickets ranging from the first to the second quantile etc. The size of circles correspond to the number of data points on each position. (C) Estimated thresholds for the ITM with 9 free threshold parameters (solid blue line), the LTM with 2 free threshold parameters (dashed red line) and the BOM with 2 free threshold parameters (dash-dotted yellow line) (D) Posterior predictive mean and 95% HDI of the LTM (dashed red line) and the BOM (dash-dotted yellow line) for Q1 to Q5, as indicated in (B). Data: solid black lines

2.2.1 *Behavioral results*

Subjects earned on average 17.1 points (SD: 4.2) in each trial (maximum points = 20), which represents a 6% loss on optimal earnings. Participants' marginal accept probabilities steadily increased as the sequence progressed (Fig. 2.1A, black line), but differed systematically from the optimal agent's accept probability (Fig. 2.1A, yellow line). On the second-to-last ($9th$) position, participants accepted the ticket only with a 28%, 95%-CI [26%, 29%], probability, whereas following the optimal policy would result in a significantly higher acceptance rate of 50%.

Overall, subjects stopped earlier than optimal. The average position at which a ticket was accepted was 4.7 (SD: 2.9), whereas an optimal agent would have stopped at an average stopping position of 5.2 (SD: 2.8). However, a closer look at Fig. 2.1A reveals that whether subjects accept too early or too late depends on the position: on earlier positions they accept options although they should continue to search, whereas, if they get to position 7, they continue searching even for options that should be accepted according to the optimal policy.

Fig. 2.1B shows the accept probabilities conditional on ticket prices, split into the first five quantile ranges $Q1$ - $Q5$ (out of a total of ten quantile ranges). Qi is defined as the range of ticket prices from the $0.ith$ to the $0.i \quad 0.1 \quad th$ quantile of the ticket price distribution. In this experiment, the ticket distribution corresponds to a Gaussian distribution with mean 180 and standard deviation of 20. Accept probabilities for $Q4$ and $Q5$ did not reach 50% at position 9, in contrast to the optimal strategy that predicts much higher acceptance probabilities at this position.

Our models did not assume any learning over trials. This assumption was supported by an analysis of performance across trials. A linear mixed model on points per trial with trial number as fixed effect and by-participant random intercepts and random slopes for trial number showed no significant effect of trial number, F 1,64.00   0.02, p 0.88.

2.2.2 *Modeling results and discussion*

First, we checked whether the key assumptions of the modeling framework were supported. We calculated, per participant and model, posterior predictive p-values (p_{pp}) that compared misfit (i.e., deviance) of the observed data with misfit of synthetic generated data from the model. For the baseline model, ITM, this analysis indicated that the absolute fit was very good, and a probabilistic threshold adequately describes participants' responses; p_{pp} .05 for only 8% of participants (Supporting Material, Fig. 2.6A). For the vast majority of participants the observed misfit was consistent with the assumptions of the ITM plus sampling variability.

The performance of the LTM was almost identical to the ITM, suggesting that the considerably more parsimonious LTM (three free parameters for LTM compared to ten for ITM) adequately describes behaviour in optimal stopping tasks. The distribution of p_{pp}-values of the LTM was almost identical to the ITM (Supporting Material, Fig. 2.6A-B). Fig. 2.1D provides qualitative evidence of the agreement between LTM and data; the LTM adequately predicts accept probabilities for each quantile at every position (see Supporting Material, Fig. 2.7 for agreement between ITM and data). Fig. 2.1C compares the recovered thresholds of ITM and LTM and shows that the ITM thresholds essentially form a straight line lying exactly on top of the LTM thresholds.

The absolute fit of the BOM is clearly worse than for ITM/LTM; p_{pp} .05 for 35% of participants (Supporting Material, Fig. 2.6C). The source for this increased misfit can be seen in Fig. 2.1D. Only for Q1 and early positions of Q4 and Q5 did the BOM provide an adequate account. Furthermore, the recovered thresholds (Fig. 2.1C) of the BOM clearly differ from the ITM in almost all positions. Results of the CoM are not shown explicitly as its performance was extremely poor. All p_{pp} 0; there was not a single posterior sample for which the observed misfit of the CoM was smaller than for synthetic data generated from the CoM. Furthermore, choices were essentially random for CoM with $_{\mathrm{CoM}}$ 0.02 [0.01, 0.06] (for the other models, 0.21).

Participants differed in their first threshold and slope parameters estimated by the LTM. However, all slope parameters are larger than 0 indicating that all participants increased the thresholds over the sequence (see also Supporting Material, Text C).

These results suggest that humans use a linear threshold when searching for the best option. In the present tests we found that the human performance is only 6% off from the performance of an optimal agent, indicating that the linear strategy performs quite well. Therefore, using linear thresholds could be an ecologically sensible adaptation to sequential choice tasks. However, it could also mean that the LTMs good performance might not generalize to new task environments, in which the linear model performs less well – an ability that would be crucial for the LTM to be a useful model of human behavior.

Search behavior in Experiment 1 indicated that people deviate from the optimal model depending on the price structure of the sequence: In trials with good options in the beginning people tended to accept them too early. However, in trials with few or no good options they continued search longer than the optimal model prescribed (Supporting Material, Fig. 2.8). Accordingly, in tasks with plenty of good options people might search less than optimal. However, in tasks in which good options are rare they might be tempted to search too long.

To find out and further predict how people will adapt to the tasks, we conducted a simulation study comparing the optimal solution with a best performing linear model (using a grid search to find the best performing parameter values for the linear model) and an empirical study manipulating the distributions of ticket prices across three conditions: (1) a left skewed distribution simulating a scarce environment, (2) a normal distribution, (3) a right skewed distribution simulating an environment with plentiful desirable alternatives. As illustrated in Supporting Material, Fig. 2.9, the simulation study showed that the optimal model predicts more search in a plentiful environment, whereas a linear model predicts more search in the scarce environment. Furthermore, the linear model predicts a stronger decline in performance in the scarce environment than the optimal model (Supporting Material, Fig. 2.9 A).

2.3 EXPERIMENT 2: DOES THE LTM GENERALIZE TO NEW ENVIRONMENTS

To show that the LTM can capture people's choice behavior across different tasks and allows us to predict when people will search too much or too little we conducted a second experiment changing the distribution of options. We manipulated the different task environments by sampling tickets from (1) a left skewed (PERT[1](40,195,200)), (2) a normal (PERT(90,140,190)) or (3) a right skewed distribution (PERT(120,125,400)), representing a scarce, a normal and a plentiful environment, respectively (Supporting Material, Fig. 2.4B-D, red lines, participant's estimate in black lines). Each participant was assigned to only one condition. The final sample included 172 participants. The procedure was identical to Experiment 1, consisting of a learning phase, where participants got acquainted with the distribution (Supporting Material, Fig. 2.5A-D), and a testing phase. In the testing phase, participants had to choose the lowest-priced ticket out of a sequence of 10 tickets with 200 trials (*Material and Methods*).

2.3.1 *Behavioral results*

Participants' performance increased from the left-skewed (scarce) environment to the right-skewed (plentiful) environment ($F(2,268) = 114$, $p < .0001$). As predicted by the best performing linear model, the loss compared to optimal performance was largest in the left-skewed condition, where only few good tickets occur (Supporting Material, Fig. 2.9A).

The average search length decreased from the left skewed scarce environment to the right skewed plentiful environment, $F(2,268) = 11.5, p < .0001$. This pattern also follows the predictions of the best performing linear model in the simulation

1 The PERT distribution is a special case of the beta distribution defined by the minimum (a), most likely (b) and maximum (c) values that a variable can take and an additional assumption that its expected value is $\frac{a + 4b + c}{6}$.

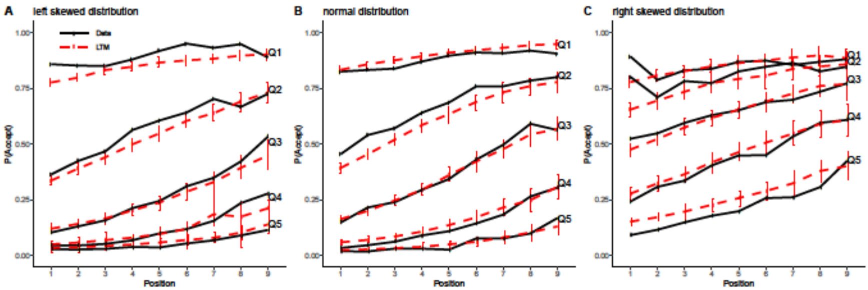

FIGURE 2.2: Results of experiment 2: Empirical data appear in black lines and the posterior predictive means of the LTM in red lines. Bars represent the 95% HDI. The different lines represent the tickets ranging in from the Q1 to Q5. Q1: Tickets in first quantile, Q2: Tickets between the first and second quantile etc. (A) Condition 1: Tickets are left skewed distributed (PERT(40,195,200)) corresponding to a scare environment. (B) Condition 2: Tickets are normally distributed (PERT(90,140,190)). (C) Condition 3: Tickets are right skewed distributed (PERT(120,125,400)) corresponding to a plentiful environment.

study but is in contrast to the optimal model's predictions (Supporting Material, Fig. 2.9B). Specifically, in the left skewed environment, where good tickets occur very rarely participants searched too long compared to an optimal agent, whereas in the environment where good tickets are abundant, participants ended their search too early compared to the optimal strategy.

2.3.2 *Modeling Results and Discussion*

Modeling results replicate the results from Experiment 1 and indicate that the LTM but not the BOM performed extremely well ($p_{\text{pp}} < .05$ for 7% to 10% of participants across the three conditions for LTM, but $p_{\text{pp}} < .05$ for 20% to 55% of participants for BOM, Supporting Material, Fig. 2.10A). The observed accept probabilities (Fig. 2.2A-C, black lines, where each line represents a ticket price within the specified quantile range) are adequately described by LTM predictions (red lines) on almost all positions and in all three environments. Moreover, the threshold parameters for the ITM are again on top of the threshold parameters estimated by the LTM in all the three environmental conditions (Supporting Material, Fig. 2.11A-C).

These results indicate that humans use a linear threshold in optimal stopping problems, independent of the distributional characters of the task. However, this does not mean that people do not adapt to the task at all. Participants are responsive to task features and adapt their first threshold and the slope to the distributional characteristics of the task within the constraints of the linear model (Supporting Material, Fig. 2.11A-C).

Experiment 1 and 2 show that the linear model reflects a robust psychological process when deciding between sequentially presented options. However, in both experiments deciders were explicitly trained on the distribution of options, something not common in real life decision making. The next experiment tests if the linear strategy can also explain choices in a realistic optimal stopping task where initial learning is omitted.

2.4 EXPERIMENT 3: LTM'S PERFORMANCE IN REALISTIC CHOICE TASK

The decision maker's goal is to buy online products at the lowest rate where prices for this product are presented sequentially. We selected commodity products from different categories (e.g food, leisure, kitchen tools) and collected for each product a set of prices from Amazon.com. Only products with approximately normal price distributions were selected for a final set of 60 products (Supporting Material, Table 2.1). In the experiment, prices were sampled from a normal distribution, with a mean and standard deviation estimated from the real prices. All participants worked on 120 trials, divided into two blocks of 60 trials. In these two blocks, the 60 products were displayed in a random order (each product was encountered twice). Participants were aware that they could see up to 10 prices in each trial, and we indicated the average price of each product on the screen to reflect that people often have an idea of familiar products' prizes and to minimize individual differences in these.

2.4.1 *Behavioral Results*

Data from 95 participants were analyzed and replicated the results from Experiments 1 and 2 (normal distribution condition). Again, participants accepted too early, on average at position 4.6 (SD: 2.9). Comparing the performance in detail to the optimal strategy showed that (Supporting Material, Fig. 2.12) participants accepted too frequently at the beginning of the sequence (i.e., too low threshold) and searched too long towards the end of the sequence (i.e., too high threshold). We again found no evidence for learning across trials (linear mixed model on points per trial with trial number as fixed effect and by-participant random intercepts and random slopes for trial number showed no significant effect of trial number $F\ \ 1,94\ \ \ \ 0.13, p\ \ \ \ 0.72$).

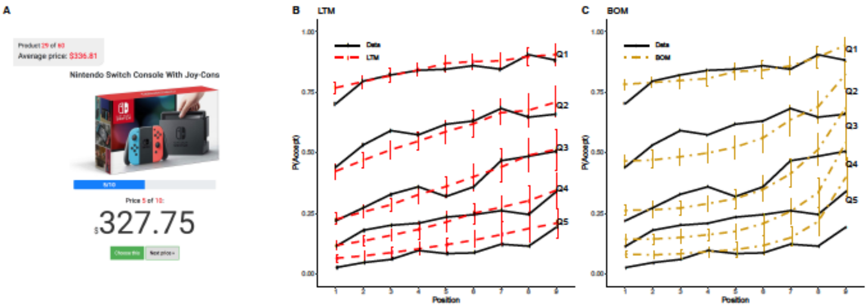

FIGURE 2.3: (A) Screenshot of the product purchasing task. (B and C) Results of experiment 3: (B) Empirical data appear in solid black lines and the posterior predictive means of the LTM in dashed red lines. (C) Empirical data appear in solid black lines and the posterior predictive means of the BOM in dashed yellow lines. Bars represent the 95% HDI. The different lines represent the product prices ranging from the first quantile to the fifth quantile. Q1: Product prices in first quantile, Q2: Product prices between the first and second quantile, Q3: Product prices ranging from second to third quantile, etc.

2.4.2 *Modeling Results*

To deal with the prices' variability we normalized all values using mean and SD prior to fitting our models. We could replicate the results from Experiment 1 and 2, despite the fact that participants did not explicitly learn the product's prices beforehand: The LTM (10% of $p_{pp} < .05$, Supporting Material, Fig. 2.14), but not the BOM (31% of $p_{pp} < .05$), was able to capture the observed accept probabilities accurately on each position and for each quantile (Fig. 2.3B&C). Furthermore, threshold parameters estimated by the LTM were very similar to threshold parameters estimated by the ITM (Supporting Material, Fig. 2.13).

2.5 DISCUSSION

In this paper, we designed a variant of an optimal stopping task that allowed us to quantitatively characterize the deviations of human behaviour from optimality. We found that humans apply a simplifying strategy, where thresholds are linearly increased over time. We implemented this assumption in a computational framework and demonstrated that this model not only provided an excellent fit to the data, it also outperformed other models found in the optimal stopping literature. Furthermore, the linear threshold assumption makes a non-trivial prediction about search length, which we confirmed experimentally: Humans stop earlier in environments with many desirable alternatives compared to scarce environments. These results contrast with the prediction from the optimal model. Finally, in a online product purchase paradigm we could show that our model generalizes to real-world sequential choice problems. Understanding how humans make sequential decisions will help quantify the conditions under which people may succeed or fail in such tasks.

But why are humans relying on a linear strategy in adapting their thresholds when an optimal policy is nonlinear? For one, our findings correspond well with recent studies demonstrating that human choice behavior in related explore-exploit paradigms is well described by a linear threshold rule (Sang et al., 2020; Song et al., 2019). But a human linearity bias seems to be more general. Indeed, a tendency to assume linear relationships has been reported in a range of domains such as function learning (Kalish et al., 2007; Lucas et al., 2015) and reasoning (Little et al., 2009; Stango et al., 2009; Wagenaar et al., 1975). Crucially, simple strategies do not necessarily perform badly. In particular in uncertain and complex environments, simple heuristics can be efficient and powerful tools if they are adapted to the structure of the environment (Gigerenzer et al., 2009; Todd, 2001). In this context, linearity could be considered as an adaptation of the human mind to its environment.

Material and Methods

2.5.0.1 *Participants*

We recruited 438 participants (272 females; age range: 18-62; $N1$ 144, $N2_{\text{left}}$ 92, $N2_{\text{normal}}$ 110, $N2_{\text{right}}$ 92, $N3$ 100 in Experiments 1, 2 and 3, respectively) on Amazon Mechanical Turk to participate in the experiments. Participants gave informed consent, and the Harvard Committee on the Use of Human Subjects approved the experiments. Participants were excluded from analysis if they accepted the first option in a trial in more than 95% of the trials. After applying these criteria, we included data from 499 participants in the subsequent analysis $\mathrm{N1}$ $129, \mathrm{N2}_{\text{left}}$ $86, \mathrm{N2}_{\text{normal}}$ $102, \mathrm{N2}_{\text{right}}$ $84, \mathrm{N3}$ 95 .

2.5.0.2 *Task*

In Exp. 1 and 2, participants performed the same online ticket shopping task that consisted of a learning and a testing phase. In the learning phase, participants experienced the distribution from which the ticket prices were drawn. In Exp. 1, the distribution from which the values were sampled was normal with 180, 20 . The procedure was as follows (Supporting Material, Fig. 2.5A-F): Participants encountered sequentially 50 ticket prices drawn from the predefined distribution. After every ten tickets, participants had to guess the average value of the tickets seen so far. After each guess, participants were told the correct response. At the end of the learning phase participants were asked to complete a histogram (by dragging the bars) for an additional 100 tickets that were drawn from the same predefined distribution. Participants received feedback by observing the correct distribution superimposed over their estimate (Goldstein et al., 2014).

In Exp. 2, we used three conditions to realize three different distributional environments, a left skewed distribution, PERT(40,195,200), a normal distribution, PERT(90,140,190), and a right skewed distribution, PERT(120,125,400). The procedure of the learning phase was identical to Exp. 1, except that we removed the section about reporting the mean for the skewed distributions (Supporting Material, Fig. 2.5B). Visual inspection of the performance in the histogram task suggested that participants learned the target distributions well (Supporting Material, Fig. 2.4).

In the second phase of Exp. 1 and 2, participants performed the ticket-shopping task. It started with a practice trial followed by 200 test trials. In each trial participants searched through a sequence of 10 ticket prices randomly drawn from the predefined distribution. For each ticket, they could decide to accept or reject it at their own speed. People were aware that they could see up to 10 tickets in each trial and they were always informed about the actual position and the number of remaining tickets (Supporting Material, Fig. 2.5E). It was not possible to go back to an earlier option after it was initially declined. If they reached the last (10^{th}) ticket they were forced to accept this ticket. When participants accepted the ticket, they received explicit feedback about how much they could have saved by choosing the lowest-priced ticket in the sequence (Supporting Material, Fig. 2.5F).

Participants were paid according to their performance. In each of the 200 trials there was a maximum of 20 points to earn. The participants received the maximum number of 20 points if they chose the lowest-priced ticket and 0 points for the worst ticket in the sequence. The payoff for a ticket that lied between the lowest-priced and the highest-priced was calculated proportional to the distance to the lowest-priced ticket in the sequence. The exact calculation for the points in each trial i was as follows:

$$points_i \quad \frac{20 \quad ticket_{max} \quad ticket_{chosen}}{ticket_{max} \quad ticket_{min}}, \tag{2.4}$$

where $ticket_{max}$ represents the most expensive ticket in the sequence and $ticket_{min}$ the cheapest ticket in the sequence. Participants received a base payment of \$4 and earned between \$0 and \$4 additionally depending on their performance.

In Exp. 3, participants performed an online product shopping task that started with a practice trial followed by 120 test trials divided into two blocks containing the same sixty products. In each trial, they encountered a product and searched trough a sequence of ten prices. Prices were randomly drawn from a normal distribution with a mean and standard deviation estimated from realistic prices collected from Amazon.com. Participants received a base payment of \$2 and a performance contingent bonus between \$0 and \$4.

Acknowledgements

We thank Michael Lee and Peter Todd for helpful reviews and are grateful to Vassilios Kaxiras for helping with data collection. CB was supported by Swiss National Science Foundation (SNSF) (168889). HS was supported by SNSF (100014_179121). SJG was supported by the Office of Naval Research (N00014-17-1-2984). BvH was supported by SNSF (157432).

Data Availability

Data and modeling scripts are available on the Open Science Framework: https://osf.io/wqth3/

2.6 SUPPORTING MATERIAL: A LINEAR THRESHOLD MODEL FOR OPTIMAL STOPPING BEHAVIOR

Text A: Calculation of optimal thresholds

We describe the calculation of optimal thresholds applied to our scenario, where payoff is proportional to the chosen value and the goal is to find cheapest ticket price. We first derive the optimal solution mathematically based on the paper of Gilbert and Mosteller (Gilbert et al., 1966, Section 5b) and further provide a more intuitive explanation.

Let's assume that a sequence of *n* ticket prices is drawn from a standard normal distribution with density:

$$f \;\; x \qquad \frac{1}{\sqrt{\quad 2}}e^{\;\frac{1}{2}\, t\,\overline{\quad}^{\,2}} \tag{2.5}$$

and the goal is to find the lowest ticket price in this sequence.

The *optimum strategy* for this task is:

If n 1, the decision maker has to accept the ticket, therefore the threshold of the last ticket is:

$$T_1 \tag{2.6}$$

The expected price of the last ticket (P_1) is the mean ( ) of the distribution.

If n 2, the decision maker decides to keep the first option or to reject it and to go on to the second one. If he goes on, his expected ticket price is . Therefore he keeps the current one, x, if x , rejects it if x and is indifferent if x . Therefore, the expected price of the last ticket (P_1) is also the threshold for the second last option:

$$T_2 \quad P_1 \tag{2.7}$$

Than for n 2, his expected price (P_2) is:

$$P_2 \qquad {}^{P_1} f \; x \quad xdx \quad P_1 \qquad {}_{P_1} f \; x \; dx \tag{2.8}$$

The remaining terms of the sequence can be computed in a recursive manner. For each n, the decision maker accepts the ticket if it is lower than the expected price of the remaining n 1 tickets (x $P_{n \;\; 1}$) but rejects if the ticket is higher than the remaining expected price ($x \quad P_{n \quad 1}$) therefore the threshold on the n-th position (T_n) is:

$$T_n \quad P_{n \quad 1} \tag{2.9}$$

Accordingly the expected price (P_n) is :

$$P_n \qquad {}^{P_{n \quad 1}} f \; x \quad xdx \quad P_{n \quad 1} \qquad {}_{P_{n \quad 1}} f \; x \; dx \tag{2.10}$$

Intuitive explanation

The optimal thresholds T_i for maximising the payoff is calculated working backward from the last ticket price: The threshold of the final item (T_1) is , because the rules of the task stipulate that the final item must be accepted if no earlier item has been chosen. The thresholds for the previous items are determined by working backward from the final item, using conditional expectations. First, we calculate the expected value of the final item (P_1). For the last item, this is the expectation of the overall probability distribution from which the options are sampled. Therefore, to maximize expected reward on the second last position, one's policy should be to accept a particular option if it is better (in our case smaller) than the expected reward if one continues under the optimal policy. The second-to-last item should be accepted if its value is smaller than the expected value of the final item. This means that the threshold of the second-to-last item (T_2) is the expected value of the last item (P_1).

The expected value of the second-to-last item (P_2) is the expected value of the part of the probability distribution that is better (in our case smaller) than the threshold (T_2) for the second-to-last item. The probability of this expected value is the area under the probability distribution that is better than this threshold. The overall expected reward at the second-to-last position (P_2) (and therefore the threshold for the third-to-last item (T_3)) is calculated as follows: we multiply the expected value for the second-to-last item with its probability plus the expected value of the last item multiplied with its probability (which is equal to 1 minus the probability of the second-to-last item). The remaining thresholds are calculated in the same way.

Text B: Modelling

Models were implemented in a *hierarchical-Bayesian statistical framework* using JAGS (Lee et al., 2014; Plummer et al., 2003). In a Bayesian framework, information regarding model parameters is represented by probability distributions. The data is used to update *prior distributions* resulting in *posterior distributions*, which were used for inference. A hierarchical implementation allows us to fit data on the individual trial-level, while simultaneously taking into account information shared across participants via group-level distributions. Results reported in the main manuscript are based on the group-level posterior distributions, unless noted otherwise.

Fitting involved running four independent chains, each with 2000 samples drawn from the posterior distribution, with a burn-in period of 100 samples. Chain con-

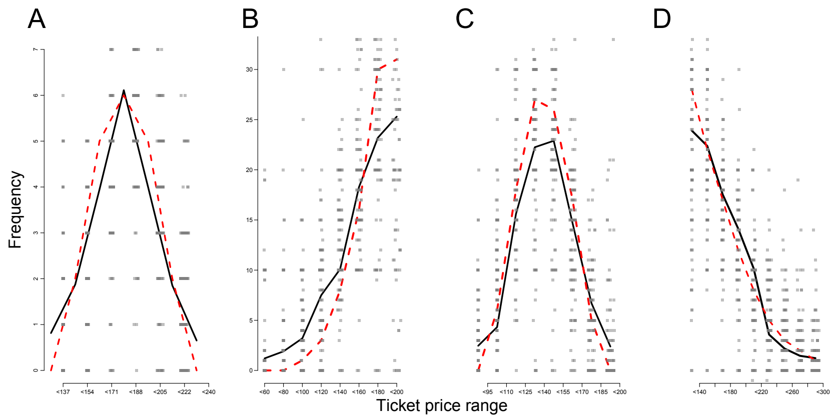

FIGURE 2.4: A-D: Results of the distribution learning phase: Participants' aggregated responses in the histogram task (details in *Methods*). Empirical data appear in black lines and the predefined distribution in red dashed lines and generally show good agreement. A: Experiment 1: Predefined distribution is a normal distribution, B-D: Experiment 2. B: Condition 1: Predefined distribution is a left skewed distribution. C: Condition 2: Predefined distribution is a normal distribution. D: Condition 3: Predefined distribution is a right skewed distribution.

vergence was monitored via the calculation of Gelman-Rubin statistics on the four chains, autocorrelation plots, and visual inspection of the chains.

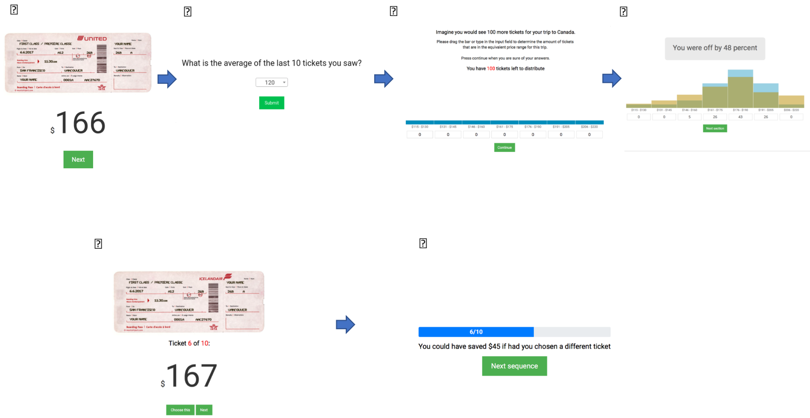

FIGURE 2.5: A-F: Screen shots of online experiment. A-D: learning phase, E-F: testing phase. A: Sequential presentation of ticket values sampled from predefined distribution. B: After each 10 tickets, participants are asked to estimate the average of the tickets just seen (this section was removed for the left and right skewed distribution in experiment 2). C: At the end of the learning phase, participants have to predict how a future sample from the same predefined population might look, where they essentially had to draw a histogram using this interface. D: Feedback was provided by superimposing the correct distribution over their estimate. E: Testing phase: In every trial participants encounter ten tickets sequentially and have to decide to accept it or to continue. Each ticket indicates the ticket's actual position in the sequence. F: Feedback was provided about how much they could have saved if they had chosen the cheapest ticket in the sequence.

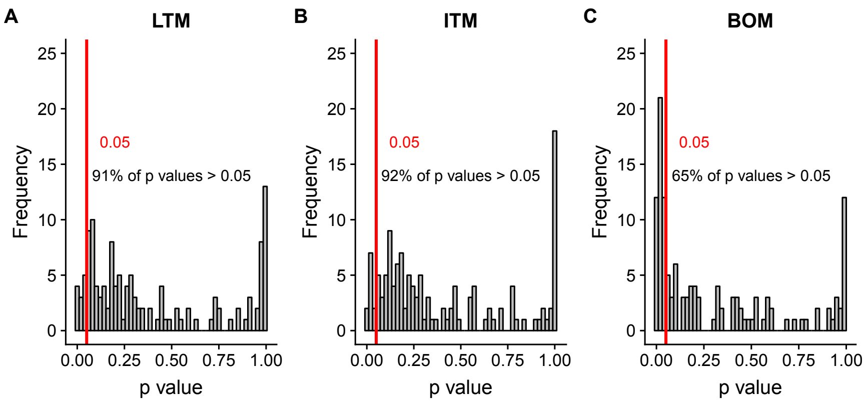

FIGURE 2.6: Results of experiment 1: Individual posterior predictive p values for the LTM (A) the ITM (B) and the BOM (C) for each individual.

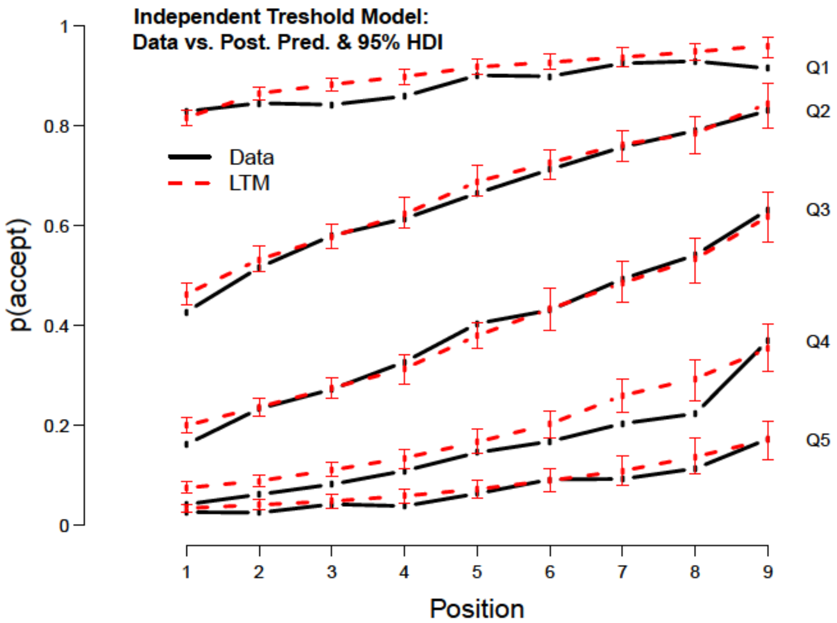

FIGURE 2.7: Results of experiment 1. Empirical data appear in black lines and the posterior predictive means of the ITM in red dashed lines. Bars represent the 95% HDI. The different lines represent the tickets ranging in from the first quantile to the fifth quantile, from a total of ten quantiles. Q1: Ticket prices ranging in first quantile, Q2: Ticket prices ranging between the first and second quantile etc.

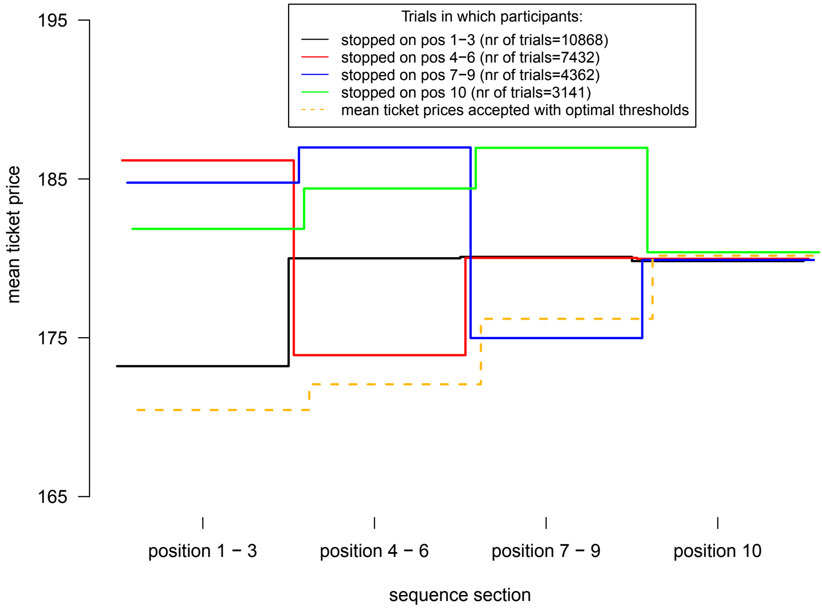

FIGURE 2.8: Illustration of the structure of prices for trials in which participants either accepted in the beginning (black line), in the middle (red line) or at the end of the sequence (blue line). Price structure in trials in which participants reached the last position are shown in green. The sequences of 10 ticket prices each for the 200 trials were generated and stored in the beginning of the experiment, ensuring that we could analyse all prices in each sequence, regardless of the stopping position of the participants. The yellow dashed line shows the mean of the accepted ticket prices when using the optimal threshold in each of the respective section. The black line shows the price structure of trials that were accepted in the beginning of the sequence, indicating that sequences that included low prices in the beginning were more likely to be accepted than the optimal threshold would prescribe (black line vs yellow line). However the blue line shows the price structure of trials that were accepted in the later part of the sequence, indicating higher prices on position 1 - 6. In these trials, participants continued search longer than the optimal model prescribed (participant's mean accepted ticket price is lower than the the optimal threshold's mean accepted ticket price, blue line vs yellow line)

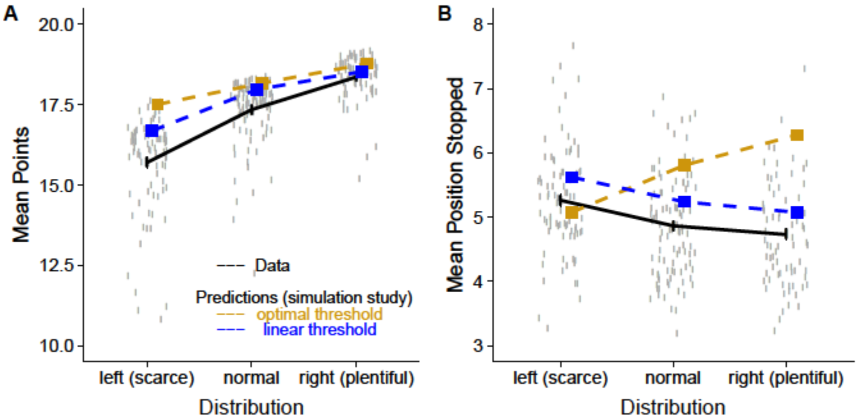

FIGURE 2.9: (A) Average performance (in points/trial) vs distributional structure of the task. Data: black line (grey dots: individual data points), performance when using optimal thresholds: yellow line, performance when using best performing linear thresholds: blue line. (B) Average search length vs distributional structure. Data: black line, individual data points: grey dots, optimal thresholds: yellow line, best-performing linear thresholds: blue line. Note that model predictions are based solely on distributional characteristics of the environments and not on model fit on human data.

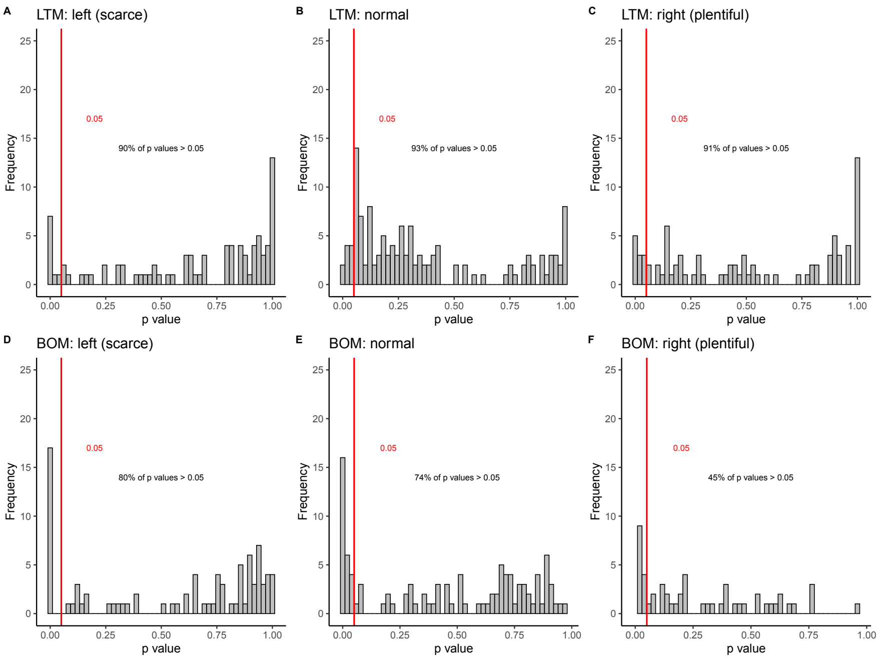

FIGURE 2.10: Individual posterior predictive p values for the LTM (A-C) and for the BOM (D-F) for each condition and each individual. A and D: Condition 1, scarce environment; B and E: Condition 2; C and F: Condition 3, plentiful environment

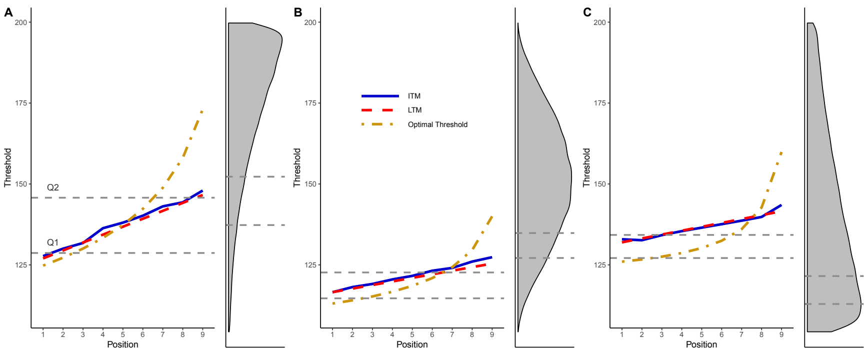

FIGURE 2.11: Experiment 2: (A) Left skewed distribution (B) Normal distribution (C) Right skewed distribution. Estimated threshold parameters from the ITM (solid blue line) and the LTM (dashed red line). The optimal threshold (no model fit) is shown as yellow dotted dashed line. The grey dashed horizontal lines indicate the first (Q1) and second (Q2) quantile of the respective distribution.

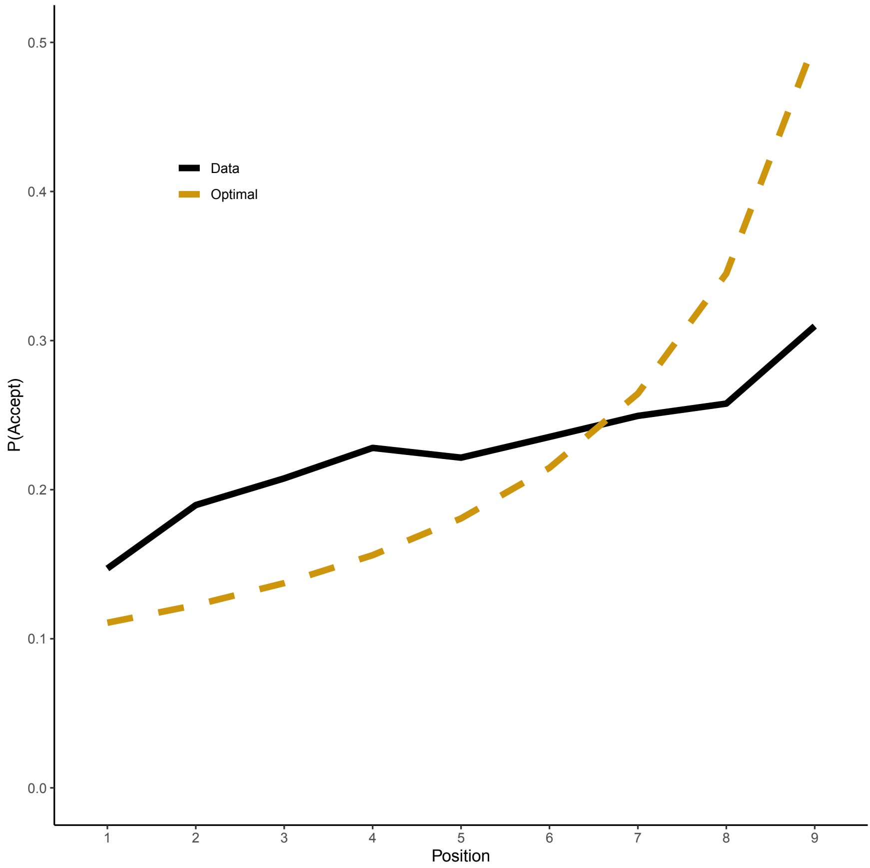

FIGURE 2.12: Experiment 3 (realistic products): Probability to accept a product price depending on the position across all prices. The dark line represents participant's frequency to accept, the dashed yellow line the optimal agent‘s probability to accept.

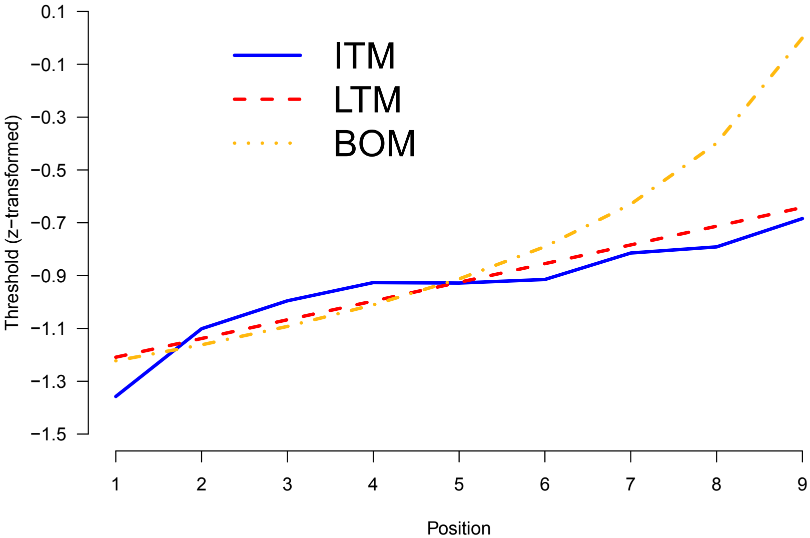

FIGURE 2.13: Experiment 3: Estimated thresholds for the ITM with 9 free threshold parameters (solid blue line), the LTM with 2 free threshold parameters (dashed red line) and the BOM with 2 free threshold parameters (dash-dotted yellow line). Product prices differed for each product, in order to make them comparable we transformed them to z-scores and calculated the thresholds a the z-scale.

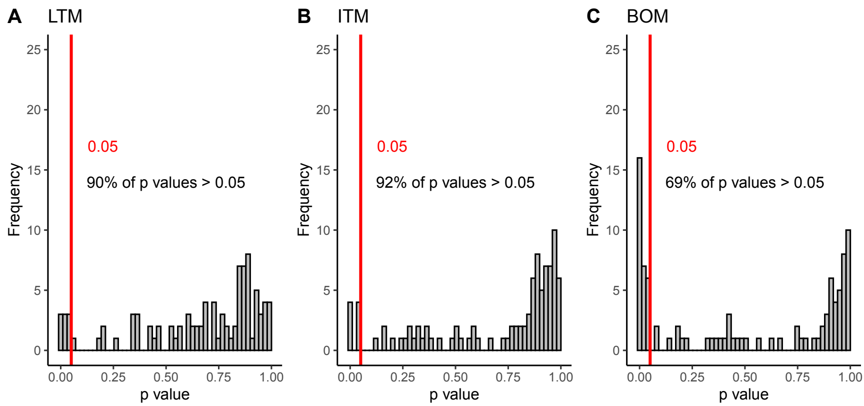

FIGURE 2.14: Results of experiment 3. Individual posterior predictive p values for the LTM (A) the ITM (B) and the BOM (C) for each individual.

Product	mean	sd
1. Philips Electric Toothbrush	62.5	3.3
2. Nintendo Switch Console With Joy-Cons	330.	26.1
3. Nintendo Switch Pro Controller	83.5	11.6
4. Adidas Men’s Stan Smith Shoes	91.1	8.5
5. Hardcover Lord of the Rings Boxed Set	53.9	7.3
6. JBL Flip 4 Portable Speaker	87.7	9.7
7. Game of Thrones: The Complete Sixth Season	32	6.1
8. GoPro Waterproof Digital Action Camera	311.	50.4
9. Crocs Adult Unisex	37.6	2.8
10. Ray-Ban Men’s Wayfarer Sunglasses	147.	13.8
11. Fjallraven Kanken Backpack	83.9	8.5
12. Nespresso Inissia Espresso Machine	146.	11.5
13. Monster Energy Drink, Zero Ultra - 24 pack	36.4	1.1
14. Black+Decker Rice Cooker and Food Steamer	34.3	6.7
15. Dyson V7 Trigger Cord-Free Handheld Vacuum Cleaner	242.	34.8
16. HP 952 Cyan, Magenta Ink Cartridges, 3 Cartridges	69.6	3.8
17. Maglite LED 3-Cell D Flashlight	35	2
18. Bosch GSR18V-190B22 18-Volt 1/2 Cordless Drill/Driver Kit	128.	16.2
19. Disney Pixar Toy Story Ultimate Walking Buzz Lightyear	43.9	8.2
20. TheraBand 23025 55 cm Pro-Series Exercise Ball Slow Red	36.4	7.6
21. Nikon 8252 ACULON A211 10-22x50 Zoom Binocular	169.	13
22. Howard Leight by Honeywell Impact Sport Sound Amplification Electronic Shooting Earmuff	66	8.7
23. STIGA Pro Carbon Performance-Level Table Tennis Racket	76.3	13.4
24. Winmau Blade 5 Bristle Dartboard	79.4	6.5
25. Coleman Sundome 4-Person Dome Tent	79.5	14.2
26. Coleman Camping Chair	31.5	5.8
27. Fuzion X-3 Pro Scooter (2018 Gold)	96.9	15.7
28. Roller Derby Women’s V-Tech 500 Button Adjustable Inline, Mint	48.4	3.2
29. Quality Suites Orlando Lake Buena Vista - Orlando, 4.6 miles to Walt Disney World Resort – 7 nights	95.4	12
30. NBA – Ticket - LOS ANGELES LAKERS VS GOLDEN STATE WARRIORS - Staples Center - Los Angeles, United States	122.	11.7
31. Hardcover Novel: Where The Crawdads Sing	21.3	3.6
32. Apple EarPods with Lightning Connector	19.1	2.4
33. BIC Soleil Women’s Disposable Razor	9.7	2.1
34. Listerine Total Care r	14.1	3.8
35. Harry Potter and The Sorcerer’s Stone (Hardcover)	25.3	5.2
36. Queen: The Platinum Collection	22.4	7
37. Fujifilm INSTAX Mini 2 Packs	16	1.7
38. Haribo Gummi Candy, 5 Pound Bag	15.3	0.9
39. Duracell Alkaline Batteries AA, 48 Count	24.1	4.3
40. Charmin Ultra Soft 2 - 12 rolls	24.3	3.1
41. Twister	20.5	2.3
42. La Roche-Posay Anthelios Ultra Light Sunscreen Fluid, SPF 60 - 1.7 fl oz bottle	28.7	3.9
43. Vitafusion MultiVites Gummy Vitamins for Adults, Assorted - 150 count	12.1	2.9
44. Extra Strength Bayer Aspirin 500mg Coated Tablets, 100ct	12.5	0.9
45. Quest Nutrition Quest Bar, Chocolate Chip Cookie Dough - 12 bars, 2.12 oz each	26.6	2.3
46. Barbie Dreamtopia Mermaid Doll 3	16.3	4.2
47. Monopoly Board Game	18.5	2.9
48. CamelBak Eddy Water Bottle, Dragonfruit, 0.75 L	12.8	2.3
49. Bear Grylls Fire Starter Gerber	17.9	3.9
50. Victorinox Swiss Army Classic SD Pocket Knife	16.9	2.4
51. Wilson NFL Super Grip Official Football	30.4	5
52. Biofreeze Professional Pain Relieving 360 Spray 4 oz	14.3	2.9
53. Bell Sports Bicycle Combination Cable Lock 5’ Watchdog 100, Black	11.2	2.5
54. Frogg Toggs FTP1714-12 Action Poncho	14.4	1.9
55. Speedo Vanquisher 2.0 Mirrored Goggle Silver	20.5	3.9
56. Intex Unicorn Inflatable Ride on Pool Float	20.6	6.2
57. Rain-X 5079280-2 Latitude 2-in-1 Water Repellency Wiper Blade - 24-inches	23.8	3.3
58. The Easy 5-Ingredient Ketogenic Diet Cookbook	12.8	1.9
59. Speck Apple iPhone XR Presidio Case	30.1	2.4
60. RoomMates Lisa Audit Butterfly Quote Peel and Stick Wall Decals	17.9	4.9

TABLE 2.1: Products with mean and standard deviation of prices

Text C: Individual differences

2.6.0.1 *Experiment 1*

Figure S12 A shows the posterior individual-level threshold parameters (transparent red lines) and posterior group-level threshold parameters (solid red line) of the Linear Threshold Model (LTM). Participants differ in both their first threshold parameter (95% range between 144 and 168) and their slope parameters (95% between 0.3 and 2.6) but they all show the same general pattern (i.e., all participants' slope parameters are positive). Figure B and C show the individual threshold and slope parameters with respect to optimality: The x axis represent the parameter values and the y axis represents the percentage difference in performance from optimality, where negative numbers indicate worse than optimal performance. We see that there is an inverse u-shaped relationship between the first threshold and difference in performance from the optimal policy (i.e., performance for participants with a first threshold near the mean show almost optimal performance but participants with first thresholds further away from the mean in either direction do not do so), but no relationship between the slope and difference in performance from the optimal policy.

2.6.0.2 *Experiment 2*

The figures in columns A-C in Figure S13 represent the three environmental conditions: Column A correspond to the scarce environment (condition 1: ticket prices are sampled from a left skewed distribution), column B corresponds to the environment with normal distributed ticket prices (condition 2) and column C corresponds to the plentiful environment (condition 2: ticket prices are sampled from a right skewed distribution). The top row shows the posterior individual-level threshold parameters (transparent red lines) and posterior group-level threshold parameters (solid red line) of the Linear Threshold Model (LTM) for each condition. Individual threshold and slope parameters vary between participants in each condition but the group-level variance (which capture inter-individual variability) is higher in condition 1 than in condition 3 ($\frac{cond1}{cond3}$ 1) for the first threshold parameter (1.24 CI: [0.93,1.66]) and the slope parameter (2.26 CI: [1.46,3.62]). Further, all the slope parameters are larger than 0 confirming the general trend to use a positive increasing threshold also in changing environments. Figuers in the middle and bottom row show the individual threshold and slope parameters compared to the difference from the optimal performance (in percentage, negative numbers indicate a worse performance than optimal). As in the first experiment, we find a u-shaped relationship between the first threshold and difference in performance from the optimal policy in condition 2 and 3, but not in condition 1. However we find no discernible relationship between the slope and difference in performance from the optimal policy.

2.6.0.3 *Experiment 3*

Figure S14 A shows the posterior individual-level standardized threshold parameters (transparent red lines) and posterior group-level standardized threshold parameters (solid red line) of the Linear Threshold Model (LTM). We observe from these figures that there are differences between individuals in both first thresholds and slopes. However the participants overwhelmingly follow the same general pattern (only two participants have a negative slope parameter). B and C show the individual parameters in comparison to the difference in optimal performances (in percentage).

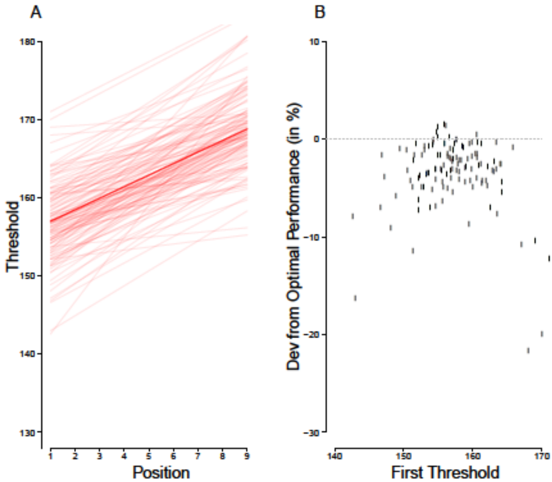

FIGURE 2.15: Individual differences in experiment 1: (A) Posterior individual-level threshold parameters (transparent red lines) and posterior group-level threshold parameters (solid red line) of the Linear Threshold Model (LTM). (B) Scatter plot of the individual threshold parameters (x-axis) and its deviation in performance from optimal strategy (y-axis) (C) Scatter plot of the individual slope parameters (x-axis) and its deviation in performance from the optimal strategy (y-axis).

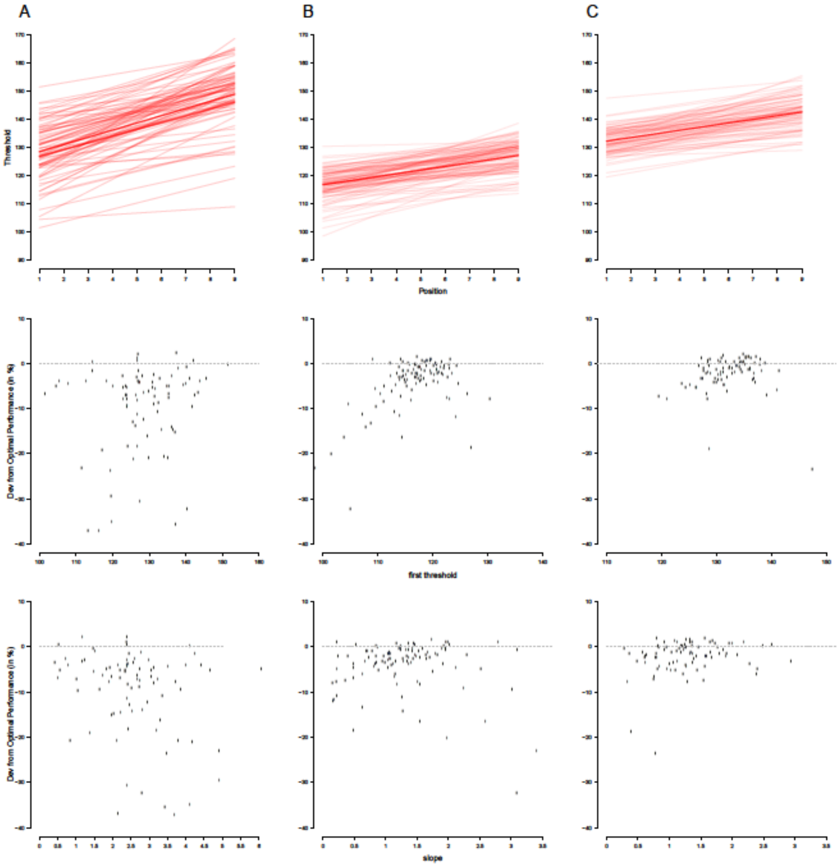

FIGURE 2.16: Experiment 2: (A-C) Top row: Posterior individual-level threshold parameters (transparent red lines) and posterior group-level threshold parameters (solid red line) of the Linear Threshold Model (LTM). (A) Condition 1, scarce environment; (B) Condition 2; (C) Condition 3, plentiful environment. Middle row: Scatter plot of the individual threshold parameters (x-axis) and its deviation in performance from optimal strategy (y-axis) for each condition. Bottom row: Scatter plot of the individual slope parameters (x-axis) and its deviation in performance from the optimal strategy (y-axis) for each condition.

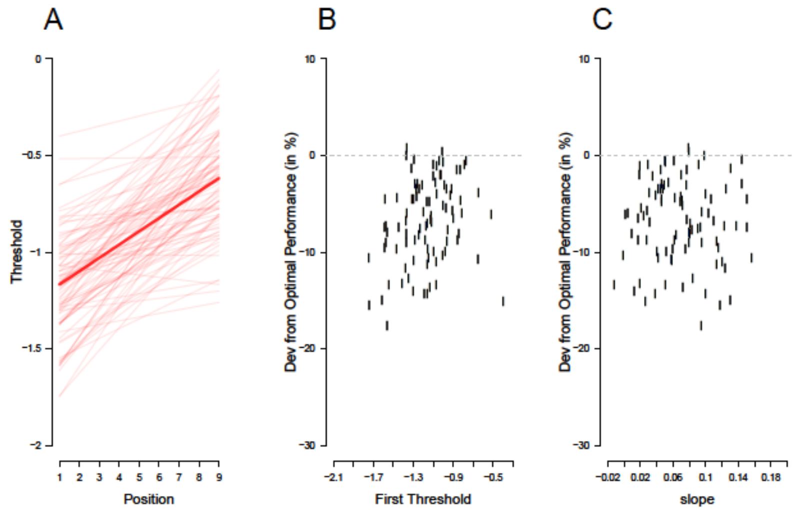

FIGURE 2.17: Experiment 3: (A) Posterior individual-level threshold parameters (transparent red lines) and posterior group-level threshold parameters (solid red line) of the Linear Threshold Model (LTM). (B) Scatter plot of the individual threshold parameters (standardized) and deviation in performance from optimal strategy (y-axis) (C) Scatter plot of the individual slope parameters (standardized) and deviation in performance from optimal strategy (y-axis).

ADAPTIVE BEHAVIOR IN OPTIMAL SEQUENTIAL SEARCH

Authors Christiane Baumann, René Schlegelmilch and Bettina von Helversen

Status submitted to Journal of Experimental Psychology: General (March 30, 2021)

ABSTRACT

Sequential decision making – a decision where available options are encountered successively - is a hallmark of our everyday life. When we search for a job or an apartment, we may decide to accept or reject it without knowing potential future options. Despite the great advances in understanding optimal stopping tasks, little is known about the adaptive behavior to changes in choice context. In this paper, we present two experiments where (1) outcome variance and (2) time horizon is modified. First we will provide empirical evidence that people adapt to both context manipulations. Secondly, we apply a recently developed threshold model of individual performance to our data that allows to separate different cognitive processes that are involved in optimal stopping behavior. Results from Study 1 show that that participants adapt perfectly to the variance of the sampling distribution and thus suggests that the value of an option is perceived relative to the options within

the sequence. Results from Study 2 suggest that participants adapt their aspiration level to time horizon in two ways: First, a short time horizon results in a more relaxed initial acceptance level leading to a higher acceptance rate. Additionally, in a short time horizon, this acceptance level is stronger adjusted during search. The studies provide insights into the underlying processes that guide adaptive behavior in optimal stopping tasks, thus providing an important step towards understanding human sequential decision making.

3.1 INTRODUCTION

To accept a safe offer or to continue searching for a better alternative is a hallmark of everyday life, whether house-hunting, selling or buying stocks, or finding a partner. In these so called optimal stopping problems, options are presented sequentially and people must choose the best option among alternatives (see Ferguson, 1989, for an early review). An important characteristic of such tasks is that there is no going back to an earlier option after it is initially declined. Therefore, the difficulty lies in the trade-off between accepting a possibly suboptimal option prematurely or rejecting it, hoping for a better one in the future. Previous research has emphasized on either studying choice behavior in comparison to normative models (see, e.g. Brickman, 1972; Corbin et al., 1975; Kahan et al., 1967; Rapoport et al., 1970) or introduced alternative choice models to characterise optimal stopping behavior (Baumann et al., 2020; Bearden et al., 2006; Cox et al., 1989; Goldstein et al., 2020; Guan et al., 2015; Hey, 1982; Lee, 2006; Lee et al., 2004; Sang et al., 2020; Seale et al., 2000; Zwick et al., 2003). These studies have shed light onto humans' simplifying strategies when decisions are made repeatedly.

However, an often neglected but important challenge when solving optimal stopping problems is to adapt one's behavior to the changes in the environment, such as dealing with outcomes of different magnitude or with limited time horizons for searching. As an example, let's assume it is Black Friday (sales for three days) and Barbara is bargain hunting for a specific video game. She explores offers from different discounters which range between $10 - $20, but the game sells fast and offers disappear quickly. At the same time, a luxury mall extends the Black Friday days over two weeks, and Barbara intends to purchase an expensive bag that is now offered for a price ranging between $700 and $900. However, price offers are changing from day to day and Barbara has to choose the right time to purchase the bag. How does Barbara adjust her choice strategy to different time and price scales?

The idea that humans are adaptive decision makers and can adjust their decision strategies to different situations is well established (Newell et al., 1972; Payne, 1982; Simon, 1990; Todd et al., 2012). Likewise, studies on optimal stopping behavior have consistently shown that people are sensitive to the environmental distributions of the values. (e.g. Corbin et al., 1975; Cox et al., 1989; Guan et al., 2018; Lee et al., 2004; Shapira et al., 1981). For example, some studies found diverging choice behavior in non-stationary environments in which presented values tended to increase or decrease over the presented sequence. (Brickman, 1972; Shapira et al., 1981). Other studies provided evidence that the skew of the options' value distribution has an effect on optimal stopping behavior (Baumann et al., 2020; Guan et al., 2018; Guan et al., 2014; Kahan et al., 1967; Rydzewska et al., 2018, e.g.) whereas plentiful environments (many good options) resulted in higher aspiration levels relative to scarce environments. Additionally, Lee at al. (2004) have shown that acceptance levels are adapted to the amount of available options, with shorter sequences leading to higher levels of acceptance (see also Guan et al., 2020; Seale et al., 1997). However, despite the consensus that people adjust their choices to different contexts in optimal stopping tasks, little is known about their adaptation strategies. Therefore, in this paper, we will investigate the robustness and reasons behind the adaptive behavior in optimal stopping problems.

Our analysis relies on the linear threshold model (LTM, Baumann et al., 2020) that allows to measure different cognitive processes that are involved in optimal stopping behavior. In particular, it assumes that human choices in optimal stopping problems can be described by (1) an aspiration or acceptance level which reflects the option value above which an option is accepted directly in the first decision (2) an adaptation rate, reflecting the change of the aspiration level with the time/number of available options running out in a linear manner and (3) choice sensitivity (or determinism), reflecting how sensitively people react to the extent by which the current option value deviates from the (adapted) aspiration level. The model thus provides a framework to measure the impact of changing task features on these mechanisms.

Our experiments include the variation of two contextual factors: (1) the outcome variance and (2) time horizon. For Study 1, participants performed an optimal stopping task where options where sampled from a distribution with either a low or a high variance. This manipulation reflects real-life optimal stopping scenarios, e.g. when bargain hunting for the best price of a tooth brush (where prices vary little) versus hunting for the best price of a luxury bag (which prices can vary largely). The question is how people adapt to such different value ranges in order to succeed. Research in simultaneous decision making has repeatedly shown that valuation is sensitive to the range of other options (e.g. Nieuwenhuis et al., 2005; Padoa-Schioppa, 2009; Rigoli et al., 2016; Tversky et al., 1993). Moreover, some researchers suggest that people value an alternative based on the relative rank within a pool of alternatives, and not by their absolute value (e.g. "normalization hypothesis", Rangel et al., 2012). Following this line of research, we would expect that variance has little effect on search performance (i.e. on accepting optimal options), but that, in order to do so, people transform their value representations into value rankings taking the variance of the alternatives into account.

Alternatively, research on risky decision making suggests that variance in the outcomes changes people's risk preferences, where higher outcome variance implies higher risk and thus results in more risk averse behavior (Genest et al., 2016; Holt et al., 2002; Markowitz, 1959; E. Weber et al., 2004). From this perspective, an optimal stopping task can be viewed as repeated risky decisions, while accepting the current option would reflect a 'safe' choice, and rejecting the current option would reflect choosing an uncertain outcome ('gamble'). If variance influences risk preference, we would expect behavioral changes under different value distributions, where higher variance would lead to enhanced risk averse behavior, translating into decreased aspiration levels in the LTM, compared to lower variance.

For Study 2, we consider the behavioral impact of the time horizon, that is, the number of options that could be explored over the sequence of choices. For instance, when time is limited, as in the mentioned example above where sales appear during three days, rejecting an option implies the possibility that only more expensive options remain. In this case, one should lower the aspiration level to ensure not to end up with a last, probably unsatisfying price offer. However, when more time is available to reach a decision, there is greater hope to find a satisfying option in the future thus the aspiration level can be raised. Indeed, studies have repeatedly shown that humans increase their general acceptance level (lowering aspiration and thus accept more) in shorter sequences, compared to longer time horizons (e.g. Lee et al., 2004; Rydzewska et al., 2018; Seale et al., 1997; Zwick et al., 2003) thus resulting in less search. These results indicate a qualitative existence of aspiration level effects, however they are limited in what they can tell us about the quantitative properties of this process (i.e., adaptation rate over subsequent rejections). That is, most research on search in optimal stopping problems has focused on data which reflect only the end product of the decision process, such as search length and accuracy. Here, we aim to identify and quantify the cognitive mechanisms that could account for adaptive process of sequential search to different time scales.

Finally, we are not only interested in how people adapt to the environment, but also whether individual differences in the decision parameters, such as the aspiration level and the adaptation rate, are stable across contexts (e.g., comparing low and high variance value distributions). Research on exploration-exploitation trade-offs has frequently emphasized that individual search might be driven by a general mechanism that affects search across different domains and tasks (Hills et al., 2008; Hills et al., 2015; Mata et al., 2015; Pirolli, 2007), however, attempts to link behavior across different exploration-exploitation tasks have been unsuccessful (e.g. Meyers et al., 2020; von Helversen et al., 2018). Our further goal thus is to uncover underlying cognitive processes that are consistent across search contexts and thus to contribute to a comprehensive understanding of search beyond optimal stopping tasks.

In the following, we first review previous research on sequential search and present three decision strategies that have been shown to account for human search behavior. We then describe our formal hypotheses and our general methodological approach, before we present each experiment separately.

3.2 THRESHOLD MODELS FOR OPTIMAL STOPPING

In a standard optimal stopping task, a decision maker encounters a series of single-valued options one at a time and must decide when to stop searching and accept a currently observed option. As an example, a decision maker (here a customer) is planing a vacation and decides to buy the plane ticket online. Ticket prices vary randomly from day to day and the customer wants to find the cheapest ticket. The customer checks the ticket price every day and decides if she wants to accept or reject the ticket, without having the option to go back in time to a previously rejected offer. Search time is limited by her vacation schedule (i.e., 10 days with one decision on each day) and, once accepted, the search ends.

In the classic version of the problem, sometimes also referred to as the "Secretary Problem", only the rank of the option relative to the ones already seen is shown. In this version of the problem, the optimal strategy is to go through the first 37% and choose the next price that is the best (has rank 1) (see Gilbert et al., 1966, for the mathematical derivation). Empirical evidence, however, suggests that, if people have such a cut-off, it is usually lower than the 37% (Kahan et al., 1967; Seale et al., 1997; von Helversen et al., 2011), thus, sub-optimally accepting too early.

The full information version of the problem differs from the classic version by showing the actual price instead of its rank. In this version, the optimal solution is derived based on the probability of finding a better price on the remaining days. From this probability, a threshold is calculated for each day (see Gilbert et al., 1966, for a mathematical derivation of the thresholds, and Supporting Material, Text 1). These optimal thresholds increase non-linearly across the time course. A threshold can be interpreted as an “acceptable” price, where anything below (when searching the minimum) is regarded “satisfactory”, anything above as “unsatisfactory”. However, studies considering the full information version of the problem have found that human choices are not fully described by assuming that they use such optimal thresholds when deciding to accept or reject an alternative (e.g. Goldstein et al., 2020; Guan et al., 2014; Lee, 2006; von Helversen et al., 2012). For instance, Guan and Lee (Guan et al., 2015) showed that despite clustering around the optimal solution, participants behavior could be better described by a *Biased Optimal Model* (BOM, Guan et al., 2015). The BOM assumes that participants’ decision thresholds deviate systematically from the optimal ones. Accordingly, this model implies that thresholds change *non-linearly* across the time course (Guan et al., 2015).

In contrast, recent studies have shown that a model assuming *linear* adaptation of thresholds better accounts for the empirical data than optimal (non-linear) threshold models (Baumann et al., 2020; Sang et al., 2020; Song et al., 2019). In a recent study we developed the *Linear Threshold Model* (LTM, Baumann et al., 2020) which assumes that humans set an aspiration level (inital threshold) at the beginning of the search process (i.e., directly affecting the first decision). Once set, the initial aspiration level is adjusted linearly with ongoing search from one option to the next (i.e., with every rejected decision, the threshold becomes more tolerant). Therefore, the strength of the adjustment and length of the sequence of choices define how far one departs from the initial aspiration level. The LTM's strength lies in its accuracy in describing human choices in optimal stopping tasks in changing environments (Baumann et al., 2020). Furthermore, its parameters can be related to psychological processes, such as an aspiration level prior to search, its adaption across time and the choice sensitivity, thus making it an appropriate model to study the adaptive processes involved in optimal stopping search.

We formally introduce the LTM, the BOM and a third model, that serves as a baseline (*Independent Threshold Model* (ITM)), using the ticket purchasing example mentioned at the beginning of this section. More formally, the customer is presented with a sequence of ticket prices $p_1, \quad , p_N$ for N sequence positions with the goal to find the minimum. If the decision maker accepts the price p_i on position i, the process terminates; otherwise, she continues to the next ticket price. When the last price p_N is reached, it must be accepted. Ticket prices vary across the sequence positions and the customer can't revert to a previously rejected offer.

According to the three models, the decision to accept or reject follows from a comparison of the current ticket price p_i with the threshold $\ _i$ (adjusted aspiration level) on the current sequence position i, and the models differ in how they calculate this adjusted threshold. The comparison between the option's value and the threshold yields an acceptance probability $\ _i$ based on a sigmoid choice function defined by

$$_i \quad , p_i, \ _i \qquad \frac{1}{1 \quad \exp \quad p_i \quad _i}. \tag{3.1}$$

The parameter (sensitivity) governs how deterministic the decision threshold is. Small values of reflect stochasticity in decisions, whereas the policy approaches determinism with larger values. In other words, large indicate that a decision maker applies the decision threshold very consistently, while low values indicate sporadic violations (e.g. either strategically to explore remaining options once in a while, or due to less precise cognitive processes during the evaluation).

The LTM postulates that the decision threshold changes over time in a linear fashion. Formally, it is defined by the first threshold $_0$ and the incremental value which is added to $_i$ whenever a price was rejected.

$$ {}_{i\;1} \qquad {}_{0} \qquad i, \tag{3.2} $$

Formally, an increasing $_i$ over time (i.e., positive values of , in the current example) means that the decision maker becomes more likely to accept more expensive plane tickets compared to the the previous sequence positions. In sum, this model entails three adjustable parameters, the first threshold $_0$, the increment of the threshold and the choice sensitivity . The corresponding psychological interpretations are listed in Table 3.1.

Parameter	Cognitive function
$_0$	Aspiration level before search
	Adaptation rate during search
	Sensitivity to differences between aspiration level and actual price

TABLE 3.1: LTM parameters.

In contrast to the LTM, the BOM (Guan et al., 2015) assumes that people use the optimal thresholds but diverge from them in a systematic way. Unlike the LTM, the BOM predicts non–linear increasing decision threshold over time. The optimal thresholds $_{i}$ for each position *i* are derived by determining the expected reward of the remaining options (derivation in Gilbert et al., 1966, Section 5b, and in Supporting Material Text 1). The model entails two parameters, a systematic bias that reflects how much above or or below their threshold is from optimal. Additionally, the thresholds depend on that determines how much the bias increases or decreases as the sequence progresses (Supporting Material Text 2 for a formal description).

The *Independent Threshold Model* (ITM) is identical to the LTM, but freely estimates $N \;\; 1$ independent threshold parameters $_{1}, \ldots, _{N \;\; 1}$, one for each position in the sequence. The thresholds can take any value across positions and therefore provides an upper limit for how well any threshold model can describe a person's decision given the assumption of a probabilistic threshold.

In following paragraph, we formulate hypotheses about choice behavior adaption to (1) variance and (2) time horizon and their theoretical implications regarding the cognitive parameters described by the LTM.

3.2.1 *Effects of variance*

In the first study, we investigate the effect of outcome variance on search behavior and decision thresholds. In two tasks, values are drawn from the same distribution with either a low or high variance, with a constant mean.

Several studies investigated how value is represented psychologically suggesting that the perceived value of a reward strongly depends on the context in which it is evaluated (Nieuwenhuis et al., 2005; Rigoli et al., 2016; Tversky et al., 1993). Accordingly, the normalization hypothesis (Rangel et al., 2012; Stewart et al., 2006) postulates that participants evaluate the prices upon their relative rather than their absolute value within the price generating distribution. The relative position of the value is thus calculated by normalizing it to the distribution of alternatives that will be encountered. Therefore, the same price in the low and the high variance environment would correspond to a higher value in the low variance environment since its percentile rank within its price distribution is lower. Furthermore, two prices that correspond to the same percentile rank in their sampling distribution would be valued equally (see also Bavard et al., 2018; Kahan et al., 1967; Klein et al., 2017; Louie et al., 2012). Under this perspective, we would not expect variance to alter search length nor performance.

This idea further implicates that the decision thresholds in low- and high-variance contexts differ with regards to the absolute scale, such that the thresholds are proportional to the variance of the value distributions. However, this difference should disappear when the value distributions are normalized to p_n 0, 1 . Figure 3.1 A1–A2 depicts the corresponding hypothesis when thresholds would be measured on an absolute-scale value distribution (A1), or on a normalized scale (i.e., standard-normal scale; A2). The normalization hypothesis implies, that there is no difference in both search length and performance between low and high variance search environments, and that a corresponding threshold adaptation should be captured on the absolute scale, but not on a normalized scale.

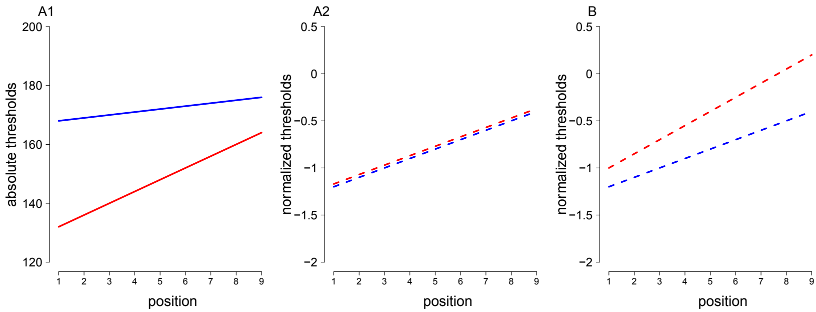

FIGURE 3.1: Thresholds predicted for low (blue line) and high (red line) variance environment. A1-A2: Normalization hypothesis: Thresholds differ on an absolute scale (A1) but are identical on a normalized scale (A2, when =0 and =1). B: Variance-risk hypothesis: High variance environment implies higher risk, therefore people accept earlier compared to low variance environment. Consequently, normalized thresholds would be higher in high variance environments than in low variance environments.

Alternatively, an optimal stopping task can be considered as a repeated risky decision between a safe (to accept) and an uncertain option (to reject) (Cox et al., 1989). Under this perspective, higher outcome variance in the gamble may lead to higher risk aversion, as has been shown in the risky decision making literature (Genest et al., 2016; Holt et al., 2002; E. Weber et al., 2004). As a consequence, acceptance rates would be increased in high variance environments, resulting in shorter search length and lower performance. Adaptive behavior to outcome variance would thus be reflected in higher decision thresholds ($_0$) and potentially also higher adaptation rate ( ), when expressed on a normalized scale (see Fig. 3.1B).

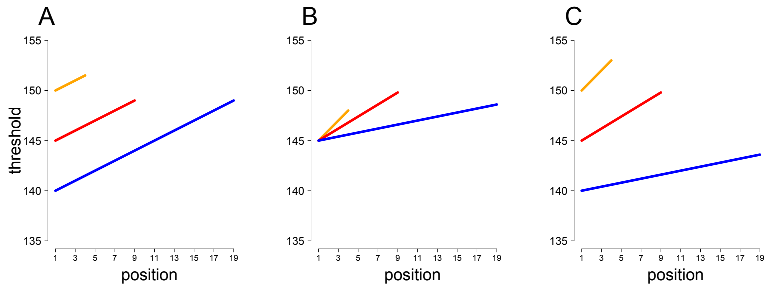

FIGURE 3.2: Hypotheses about the impact of time horizon on decision thresholds ($_0$: initial threshold, : adjustment of threshold across position. A: Time horizon affects the initial threshold $_0$). B: Time horizon affects adjustment parameter. C: Time horizon affects both $_0$ and

3.2.2 *Effects of time horizon*

In the second study, we investigated the question whether the time horizon (i.e., how many choices could be made until the ultimate one), affects choice behavior and which cognitive processes are involved. Previous studies observed that in longer time horizons, humans search more (in absolute terms) (Seale et al., 2000) and showed that decision thresholds are lower in longer sequences (Guan et al., 2015; Lee et al., 2004). This adaptation to longer time horizons could be explained either by a lower initial aspiration level prior to search or by a lower adjustment rate of the threshold across the sequence. To investigate the details of this phenomenon in our study, we used sequences with 5, 10 and 20 available choices (sequence positions). Thereby, we wanted to find out if time horizon affects (a) the initial threshold (see Fig. 3.2 A), (b) the change of threshold over time (Fig. 3.2 B), or (c) both (see Fig. 3.2 C).

To translate our LTM hypotheses into the predictions of search length and performance we conducted a series of model simulations illustrated in Fig. 3.3.[1] First, Fig. 3.3 A shows the predictions of performance (A1) and search length (A2) in the three conditions of time horizon (x-axes; 5 vs. 10 vs. 20) depending on different aspiration levels (colored circles; $_0$), without changing them over time (i.e., with set to 0). First, Fig. 3.3 A1 shows that, within each condition, the maximum performance predicted by the LTM is a curvy-linear function of the initial aspiration level. Intuitively speaking, when searching for the minimum, too high aspiration levels lead to accepting too many sub-optimal options and too low aspiration levels lead to rejecting too many optimal options.

Second, the connected dots represent identical threshold values across conditions (indicated by their color in Fig. 3.3). As can be seen, the threshold that yields optimal performance with a time horizon of 5 (orange circles) predicts only sub-optimal performance with a time horizon of 10 (the same holds for a comparison between 10 and 20). This means, according to the LTM, if participants seek to optimize their behaviour, relative to a time horizon of 5, they actually have to decrease their aspiration levels (less tolerant thresholds). Specifically, with the assumption that there is no adaptation across sequence positions (=0), the LTM predicts that the optimal initial thresholds for the conditions of 5, 10 and 20 are -0.4, -0.8 and -1.2 (on a normalized scale), respectively. Correspondingly, for these optimal initial thresholds, the LTM predicts that the relative search length (relative to the time horizon; A2) decreases with increasing time horizon.

1 For each sequence length (N=5, 10 and 20), an equivalent amount of 60 trials were simulated, where values were sampled from 180, 20 for a total of 100 participants, which were aggregated on measures of performance and search length. We approximated the range of the parameters to the range of values obtained in experiment 1 and used any possible combination of initial threshold ($_0$) and rate of increase ( ) to simulate choices

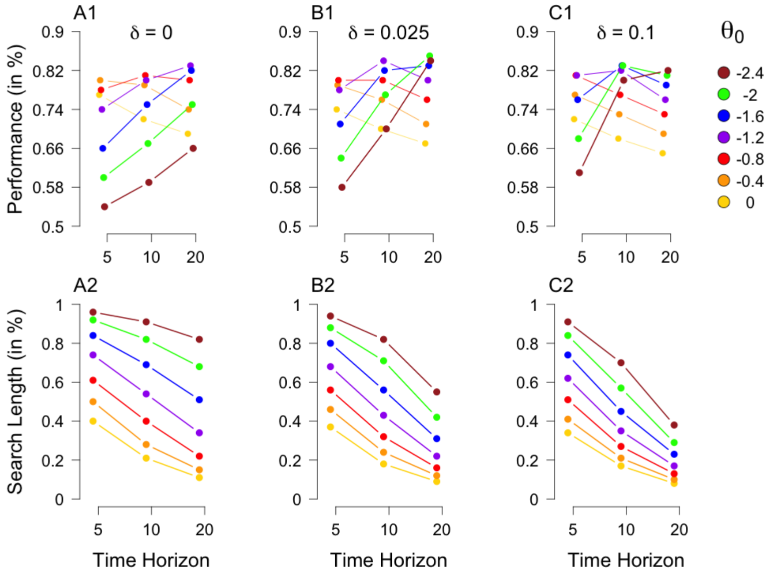

FIGURE 3.3: LTM simulation study. A1-A2: Performance (average amount of points/maximum amount of points) and search length (i/N; i = accepted position, N = time horizon) for a range of $_0$ values (see legend) and 0. Optimal thresholds differ between time horizons: N=5: $_0$=-0.4 (orange), N=10: $_0$=-0.8 (red), N=20: $_0$=-0.8 (purple). B1-B2: =0.025, $_0$ values according to legend, C1-C2: =0.1, $_0$ values according to legend. Points connected with a line represent the same $_0$ values. In (A2), (B2), and (C3), search length is calculated in relative terms (i.e., i/N; i = accepted position, N = time horizon).

Up to now, our simulation assumes that the thresholds remain fixed across the sequence. In the next step, we allow the initial threshold ($_0$) to change across positions by a constant adjustment rate . The simulation results indicate that also contributes to optimal performance in the LTM, suggesting differences between the time horizon conditions, as illustrated in Fig. 3.3 B and C. First, while 0 in Fig. 3.3 A1, introducing a small adaptation rate (0.025 in Fig. 3.3 B1) leads the previously “optimal” threshold (e.g. orange circle in A1 for time horizon = 5) to become sub-optimal. Further, whereas a small adaptation of (0.025) leads to a higher performance when N=20 (B1), a large adaption (in C1) seems to impede performance in this condition (when 0.1).

This model dynamic can be plausibly verbalized in a psychological way. That is, if participants adapt to the task, they might approach the different time horizons with different initial aspiration levels in the first place (longer time horizon = lower aspiration level $_0$, see A1). In the second place, they would regulate their adjustment of the aspiration level over time according to the time horizon with lower increases in long time horizons, because too strongly adapting the threshold over long time horizons leads to prematurely accepting sub-optimal offers. The combination of these two hypotheses schematically corresponds to Fig. 3.2C, in contrast to the alternative hypotheses, that either only the initial thresholds adapt (Fig. 3.2A) or only their change over time (Fig. 3.2B).

3.3 STUDY 1: VARIANCE

The first study investigated how outcome variance affects search behavior in the optimal stopping task. In an online purchase task, we asked participants to repeatedly perform a sequential choice task in which they choose the cheapest airplane ticket. We manipulated the variance within participants by generating ticket prices from two Gaussian distributions with high and low variance (but constant mean; i.e. 180, 10 or 180, 40).

3.3.1 *Methods*

Participants

We recruited 205 participants (78 females; age range: 19-70) on Amazon Mechanical Turk to participate in the experiment. Participants gave informed consent, and the study design and methods were approved by the ethics committee of the University of Zurich. Participant were excluded from analysis if they accepted the first option in a sequence in more than 90% of the trials with 199 participants remaining for the subsequent analysis. Participants received a fixed payment of $4 and a performance dependant bonus ranging between $0 - $4 that was calculated contingent to the choices in each trial (see following section for the exact calculation).

Procedure

The task was adapted from the study of Baumann et al. (2020)) and imitates an airline ticket-shopping scenario where people search for the cheapest price. In each trial, participants search through a sequence of ten prices. For each price, they decide to accept it or to reject it at their own pace. Participants are aware of the total number of prices in each trial and of the actual position in the sequence (see Fig. 3.4 E). It is not possible to go back to an earlier option after it is initially declined. If they reach the last price ($10th$) they are forced to choose it. When participants accept the price, they receive feedback about how much they could have saved if they had chosen the best price in the sequence. A bonus was paid proportional to the performance (see Equation 3 below).

Prior to each of the two conditions, participants learned about the respective distribution to exclude learning effects in the stopping task. During learning, participants saw 50 ticket prices drawn from the target distribution, presented sequentially in randomly ordered trials (Fig. 3.4 A). After every tenth trial, participants had to guess the overall average and the correct answer was revealed (Fig. 3.4 B). Finally, participants estimated a histogram (by dragging the bars) of the true distribution (Fig. 3.4 C), before the correct distribution was superimposed over their estimate (see Fig. 3.4 D) (method proposed in Goldstein et al., 2014). If their estimates deviated more than 20% from the correct answers, they were instructed to repeat the learning phase, with a maximum of three repetitions.

All participants completed 80 sequences in each of the high and low variance conditions. Within each condition, we sampled the ticket prices either from 180, 10 or 180, 40 . The order of the two conditions was randomized and each participant encountered a newly generated sample.

Before the task started, participants were informed that their payoff was determined by a lump sum of \$4 plus a bonus between \$0 and \$4 contingent on their performance. Each trial offered a maximum of 25 points to earn, which translated into a bonus 0.025 cents per trial. The maximum number of 25 points was awarded when the cheapest ticket was chosen and 0 points for the most expensive one. The calculation for points of tickets in between was:

$$points_j \quad \frac{25 \quad p_j^{max} \quad p_j^{chosen}}{p_j^{max} \quad p_j^{min}}, \tag{3.3}$$

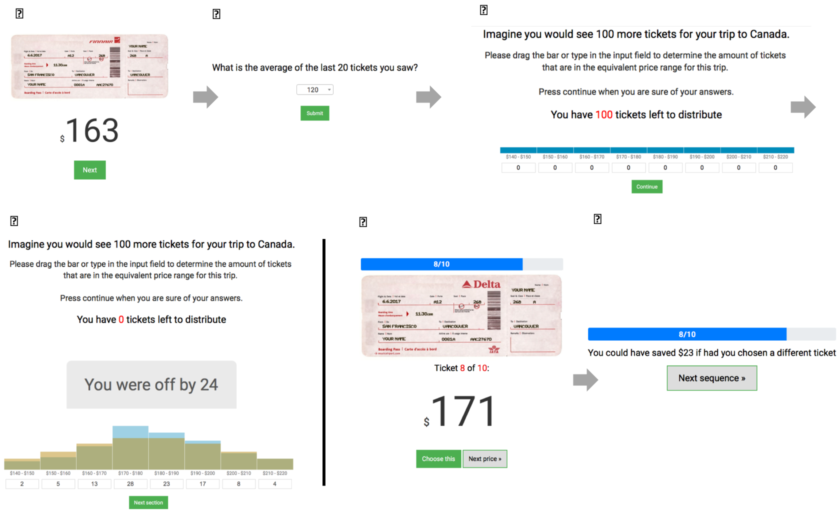

FIGURE 3.4: A-F: Screen shots of online experiment. A-D: learning phase, E-F: testing phase. A: Sequential presentation of ticket values sampled from predefined distribution. B: After each 10 tickets, participants are asked to estimate the average of the tickets just seen. C: At the end of the learning phase, participants have to predict how a future sample from the same predefined population might look, where they essentially had to draw a histogram using this interface. D: Feedback was provided by superimposing the correct distribution over their estimate. E: Testing phase: In every trial participants encounter ten tickets sequentially and have to decide to accept it or to continue. Each ticket indicates the ticket's actual position in the sequence. F: Feedback was provided about how much they could have saved if they had chosen the cheapest ticket in the sequence.

where p_j^{max} is the highest and p_j^{min} the lowest price in trial j. The final bonus was calculated by dividing the total amount of points by 1000. The average bonus earned was \$ 3.38 (min:\$ 2.83 max:\$ 3.64).

Results

3.3.2 *Behavioral results*

3.3.2.1 *Search length and performance*

We first investigated if outcome variance affects search length and overall performance. Search length is measured as the number of 'reject' decisions in a sequence plus 1. Performance was defined as the average sum of accumulated points in each trail. The visual inspection of the plotted data (Fig. 3.5 A and B) indicates that there is no difference in both measures between conditions. A Bayesian t test (R BayesFactor::ttestBF package, prior scale = medium; Morey et al., 2018) confirms that there is moderate to strong evidence against a difference between the high and low variance conditions in search length ($M^{SL}{}_{10}$ 5.03, $M^{SL}{}_{40}$ 4.99, M^{SL}_{Diff} 0.05, HDI_{95} 0.07, 0.16 , BF_{10} 0.11) and performance ($M^{Perf}{}_{10}$ 21.24, $M^{Perf}{}_{40}$ 21.36, M^{Perf}_{Diff} 0.058, HDI_{95} 0.14, 0.025 , BF_{10} 0.27). This result supports the normalization hypothesis which predicts that changing variance has no effect on search length or performance.

Fig. 3.5 C shows participant's probability (blue line: low variance, red line: high variance) to accept on each position for both variance conditions. The accept probabilities increase over position (participants become less selective) with only

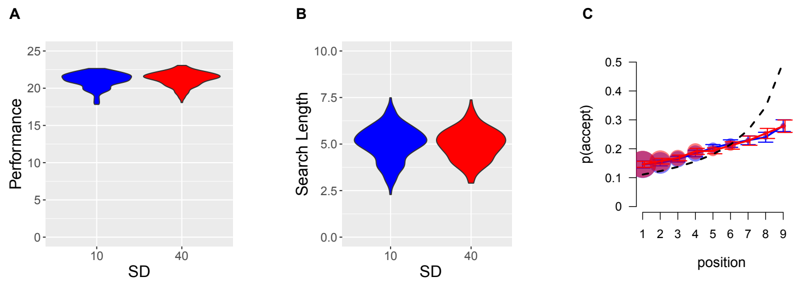

FIGURE 3.5: A: performance (maximum 25 points), B: Search length (maximum: 10 positions), blue: high variance environment, red: low variance environment. C: Probability to accept a ticket on each position across all prices. Blue line: mean and 95% CI in low variance condition, red line: mean and 95% CI in high variance condition. Black dotted line: optimal threshold rule. Points size according to number of data points (largest points: 16000 data points, smallest points: 4000 data points.

small differences between the two variance conditions. This finding further supports the normalization hypothesis, which expects no differences in accept probabilities between conditions. Furthermore, participants' accept probabilities largely deviate from predictions of the optimal threshold (dashed black line) which will be further analysed in the following modeling section.

3.3.3 *Modeling results*

Model comparison

We use a model comparison to determine which of the proposed models, the LTM, the BOM and the ITM, provides the best approximation to participants' choices. The models were implemented in a *hierarchical-Bayesian statistical framework* using JAGS (Plummer et al., 2003). The data is used to update *prior distributions* resulting in *posterior distributions*, which were used for inference. The hierarchical implementation allows to fit data on the individual trial-level, while simultaneously taking into account information shared across participants via group-level distributions. Results reported below are based on the group-level posterior distributions, unless noted otherwise. A description of all models' priors can be found in the Supporting Material (Methods).

We first tested the performance of the LTM by comparing it to a hierarchical implementation of the BOM. We compared these three models in terms of their Deviance Information Criterion (DIC) values. DIC is a model selection statistic that takes into account models' goodness of fit and penalizes them according to their flexibility (see Spiegelhalter et al., 2002). DIC differences larger than 10 are considered as strongly favouring the winning model (see Spiegelhalter et al., 2002, p.613). As it turned out, the difference between the LTM the BOM was 704 in favour of LTM. The comparison between the LTM and the more complex baseline model, the ITM, yields a difference of 37 in favour of the LTM. Figure 3.6 A displays the recovered thresholds for both the LTM and the ITM. We observe that the ITM thresholds essentially form a straight line mirroring the thresholds estimated by the LTM. These results suggest that the LTM provides a more parsimonious but equally accurate account of the data compared to the ITM which infers thresholds independently of each other.

In order to asses the accuracy of the LTM, we generated data sets simulated from the posterior distribution of the LTM parameters. Fig. 3.6 B and C shows the faithfulness of these replicates to the original data. In this figure, the accept probabilities are conditional on ticket prices, split into the first six quantile ranges $Q1$ - $Q6$ (out of a total of ten quantile ranges). Qi is defined as the range of ticket

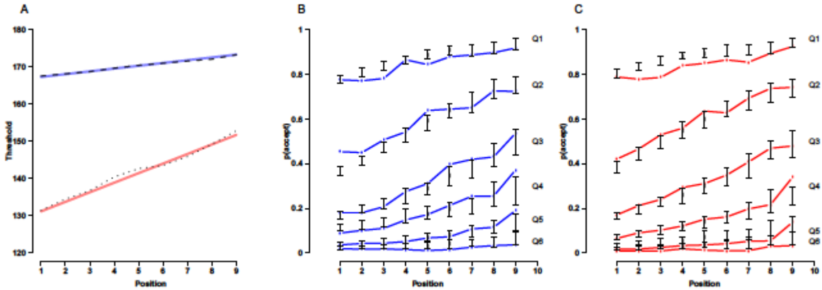

FIGURE 3.6: A: Estimated thresholds, low variance condition: LTM (blue solid line) versus ITM (black dashed line), high variance condition: LTM (red solid line) versus ITM (black dotted line) B and C: Probability to accept across positions, data versus LTM's posterior predictions. Empirical data appear in solid lines, black arrows represent the the posterior predictive means and the 95% HDI. The different lines represent prices ranging from the Q1 to Q6. Q1: Prices in first quantile, Q2: Prices between the first and second quantile etc. B: low variance condition C: high variance condition

prices from the $0.ith$ to the $(0.i-0.1)th$ quantile of the ticket price distribution. The LTM adequately predicts accept probabilities for each quantile at every position for the low and the high variance condition.

3.3.3.1 *Comparing parameter estimates between low and high variance condition*

Since the behavioral evidence supported the normalization hypothesis, the question is whether the estimated LTM thresholds differ between the low and the high variance condition. We hypothesized, that the aspiration levels (threshold θ_0) differ on an absolute, but not on a normalized outcome scale. The corresponding estimates, after the model was applied to the data, are plotted in Fig. 3.7 A (means and individual estimates). The visual inspection suggests that the estimated absolute thresholds differ substantially between the two conditions indicating that participants adjust their thresholds to variance. This finding was confirmed by comparing the group-mean parameter estimates of the initial threshold ($_0$) and its adjustment over time ( ) between the variance conditions which revealed a substantial difference between the low and high variance condition (see Table 3.2, third column).

In order to test the correctness of the normalization or the variance-risk hypothesis, we normalized prices to the same scale (=0 and =1) before fitting the model. This procedure ensures that the estimated thresholds are directly comparable on the within the distribution. The scaled thresholds yield by the group-mean parameters are plotted in Fig. 3.7 B, along with the individual-level parameters. This time, thresholds in both conditions seem to match closely. The corresponding group-level parameters are listed in Table 3.2.

The indifference of threshold estimates on the normalized scale implies that participants accept prices in both conditions within the same percentile on each position, thus providing strong support for the normalization hypothesis. Since the normalization is based on a standard-normal distribution, the resulting thresholds on each position can be interpreted akin to z-scores (e.g., z 1.1 lies above 86% of the outcomes, and z 0 in the middle; see Table 3.2). That is, at this level the model estimates a switch between accepting and rejecting an offer. However, how strongly this translates to an accept/reject decision over neighboring outcomes is co-determined by the choice sensitivity in the logistic response function, which is estimated with about 3.6 and 3.9 in each condition, reflecting a rather deterministic criterion on average.

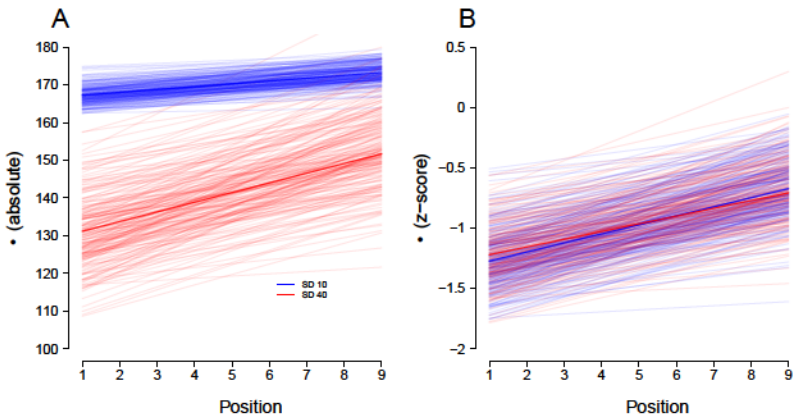

FIGURE 3.7: Individual-level threshold parameters for both variance conditions (blue: low variance, red: high variance). The two darker lines show the thresholds based on the group-level parameters. A: thresholds estimated using absolute prices B: threshold estimated using normalized prices (μ=0 and σ=1).

	Group-mean parameter estimates		
LTM parameter	$\sigma = 10$	$\sigma = 40$	Difference (σ_{10} - σ_{40})
θ_0 first threshold	168.3 [168.8, 169.2]	135.2 [134.0, 136.8]	33.4 [31.7, 35.0]
δ slope	0.7 [0.6,0.8]	2.4 [2.2,4.2]	-1.9 [-3.8, -0.3]
β choice sensitivity	0.32 [0.3,0.34]	0.1 [0.09,0.11]	0.2 [0.19, 0.24]
θ_0 scaled	-1.17 [-1.21,-1.14]	-1.12 [-1.15,-1.08]	-0.05 [-0.11,0.01]
δ scaled	0.07 [0.06,0.08]	0.06 [0.055,0.07]	0.01 [-0.01,0.2]
β scaled	3.61 [3.43,3.78]	3.89 [3.69,4.01]	-0.21 [-0.26,0.026]

TABLE 3.2: Posterior group-level mean parameters of the LTM (95% HDI in square brackets) for low and high variance condition. 3 top rows: absolute parameter values, 3 bottom rows: normalized (μ=0 and σ=1) parameter values)

3.3.3.2 *Best-fit parameters and behavioral patterns*

The goal of the following analysis is to understand the association between the cognitive parameters and the participants' performance indices, such as search length and accumulated points. A scatterplot (Fig. 3.8) demonstrates the relationship between each parameter and search length (in percent: search length/total length) and performance (in percent: points/maximum points). Visual inspection suggests that the first threshold $_0$ is negatively correlated with search length, with higher values leading to longer search. The parameters and do not appear to be related with search length. To further analyse these patterns we conducted a Bayesian linear regression analysis with search length as the dependent variable and , , and the condition (0: low variance, 1: high variance) as fixed and subjects as random effects (R BayesFactor::lmBF package, prior scale = medium; Morey et al., 2018). The analysis reveals that the best model includes all the covariates except the condition factor (see Supporting Material, Table 3.6, for the ranking of the models). As suspected from the scatterplot (Figure 3.8 A1-A3), $_0$ accounts for a large proportion of the variation of search length (R^2=0.8), leaving little room for and ($R^2 < 0.04$, see Supporting Material, Table 3.7, correlation analysis for each parameter). This result indicates that the initial threshold is mainly responsible for how long participants search, whereas the adjustment across the sequence plays just a minor role.

Visual inspection of the scatterplot between performance and $_0$ (Fig. 3.8 B1) indicates a non linear relationship, indicating that too extreme thresholds (too high or too low) impede performance, which we also highlighted in the simulation in Figure 3.3 (when Time Horzion = 10). Interestingly, the parameter reveals a strong positive association with performance, where higher values in reflect more deterministic application of the outcome thresholds (Fig. 3.8 B3)). A linear regression analysis (R BayesFactor::lmBF package, prior scale = medium; Morey et al., 2018) reveals that a model including ^{2}, and as predictors for performance provides the strongest evidence (see Supporting Material, Table 3.8, ranking of the five best models), compared to all possible models including ^{2}, , , and condition as predictors. $^2_0$ as well as mainly account for the variance in performance, with explaining 42% and 55% (respectively) of its variation (see Supporting Material Table 3.9). This result suggests that the initial threshold, which reflects the aspiration level prior to search, has a large impact on search length and performance. In particular, defining the level too low or too high in the beginning leads to poorer performance. Choice sensitivity has an additional impact on performance, where higher determinism and thus higher sensibility around the threshold leads to better performance.

3.3.3.3 *Stability of LTM parameters*

In this section we address the question of how consistent the individually fitted parameters are across both task conditions. To this goal, we correlated the obtained individual values between the two sessions (R BayesFactor::correlationBF package, prior scale = medium; Morey et al., 2018). Table 3.3 shows the results. There is very strong evidence that the cognitive parameters are moderately to strongly correlated between the two conditions. $_0$, reflecting participants' aspiration level before search, reveals the strongest correlation (=0.62), pointing to a stable mechanisms across the two tasks. Note that the first threshold ($_0$) and the adjustment rate across search ( ) are negatively correlated to some extend, showing a typical parameter dependency

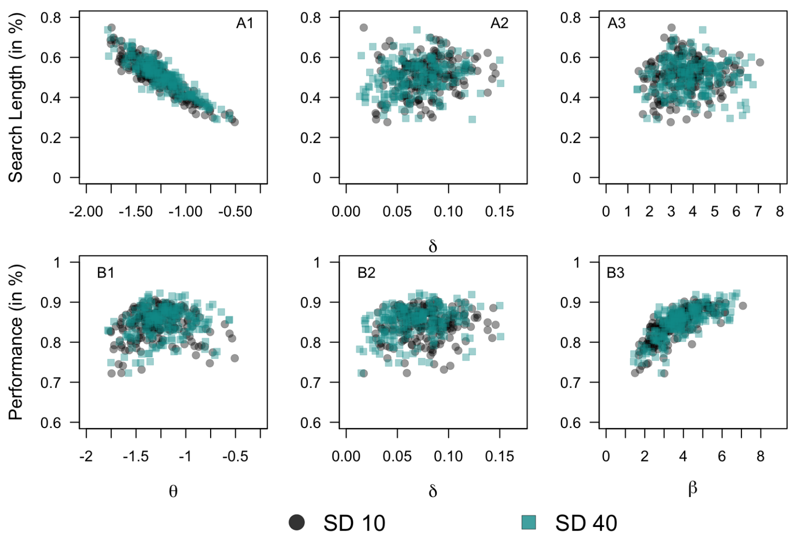

FIGURE 3.8: A1-A3: Scatterplots of search length (in %, search length/total sequence length) vs parameters ($_{0}$, and) B1-B3: Performance (in %, points/maximum points) vs parameters. Black: low variance, green: high variance

as in all regression-like (intercept plus slope) models. That is, if the initial threshold is very high in the first place, then the ceiling is almost reached and there is no point in further increases, and starting with a very low threshold, naturally, allows for the strongest increase. Importantly, all other correlations meander around 0, suggesting an independent contribution to predictions.

3.3.4 *Discussion Study 1*

In line with previous literature (Lee et al., 2004; von Helversen et al., 2012; Zwick et al., 2003), we found with the LTM that the first threshold (initial aspiration

	10	40	10	40	10
10	1				
40	**0.62**[0.59,0.65]	1			
10	**-0.35**[-0.47,-0.23]	**-0.26**[-0.38,-0.13]	1		
40	-0.06[-0.12,0.07]	**-0.34**[-0.45,-0.21]	**0.53**[0.41,0.61]	1	
10	0.08[-0.05,0.21]	0.01[-0.13,0.15]	0.08 [-0.05,0.21]	-0.13[-0.26,0.00]	1
40	0.23[0.11,0.36]	0.19[0.06,0.32]	0.03[-0.10,0.17]	0.19[0.05,0.32]	**0.56** [0.45,0.64]

TABLE 3.3: Study 1: Correlations between parameters with 95% HDI (square brackets, bold numbers: BF > 150)

level) accounts for a large portion of variation in search length (about 80%) while performance shows a curvilinear association with the first threshold parameter, indicating that setting the threshold to low or to high in the beginning of the search results in lower performance. Our analyses add on these findings by showing that the adjustment of the acceptance level during search has minor impact on search length and a negligible influence on performance. However, the parameter, that represents the choice sensitivity between the aspiration level and the actual price, accounts up to 55% of the performance' variation. Therefore, the choice sensitivity is highly related to performance, with high values (and thus more determinism) leading to a higher performance.

The model parameters consistently correlated between the low and the high variance conditions, with ranging between 0.53 and 0.62. While the magnitude of the correlations are lower than usually reported retest reliability measures (e.g.

Balloon Analogue Risk Task (BART), Lejuez et al., 2002, =0.70), one needs to keep in mind that participants encountered two different versions of the optimal stopping task. Thus, overall the results suggests that the processes captured in the linear threshold model reflect reliable individual differences.

3.4 STUDY 2: TIME HORIZON

In the second study, we focused on the influence of the time horizon on adaptive choice behavior in optimal stopping tasks, by manipulating the sequence lengths. That is, sequences contained either 5, 10 or 20 options. Otherwise, the task was equivalent to that in Study 1, where participants were conducting an online ticket-shopping task with the goal to find the cheapest price.

3.4.1 *Methods*

3.4.1.1 *Participants*

We recruited 70 participants (20 females; age range: 25 - 60) on Amazon Mechanical Turk to participate in the three-conditions within-subjects experiment. We selected a sample size equivalent to an equal study of Baumann et al. (2020). Participants gave informed consent, and the University of Zurich Committee on the Use of Human Subjects approved the experiments. Participants were excluded from analysis if they accepted the first option in a trial in more than 95% of the trials leaving 68 participants in the subsequent analysis. They received a fixed payment of \$2 and a performance dependant bonus ranging between \$0 – \$4 (calculation in the next section).

3.4.1.2 *Procedure*

The second experiment employs the same task as in Study 1. The sequence length was varied, offering either 5, 10 or 20 values. All prices were generated from the same distribution throughout the experiment (180, 20). All participants worked on a total of 180 trials, split into two parts of 90 trials. Both parts consisted of 3 blocks with 30 trials, where trials within the block where either of length N = 5, 10 or 20. The order of the blocks was randomized.

Participants were paid according to their performance. In each of the 180 trials there was a maximum of 33 points to earn, which corresponds to 0.02 cents per trial. The participants received the maximum number of 33 points if they chose the lowest (and therefore 0.02 cents) and 0 points for the highest price in the sequence (and consequently 0 cents). The payoff function was:

$$points_i \quad \frac{33 \quad p_{max} \quad p_{chosen}}{p_{max} \quad p_{min}}, \tag{3.4}$$

where p_{max} represents the highest price and p_{min} the highest price in sequence i. Participants received a base payment of \$4 and earned between \$0 and \$4 additionally depending on their performance.

Results

3.4.2 *Behavioral results*

3.4.2.1 *Search length and Performance*

To examine the effect of time horizon on the participants' performance indices, we compared search length and performance (calculated as in Study 1) between the three search length conditions (N=5, 10 and 20). Participant's relative search length (search length/total sequence length) is shown in (Fig. 3.9 A, black) for each condition. We observe that participants search length decreases in relative terms with increasing time horizon: When N= 5, participants search on average 58% (95% CI (between 0 and 1)=[0.57, 0.62], absolute search length: M: 2.9, SD: 0.49). When N=10, average search length is reduced to 48.6% (95% CI=[0.46,0.51], M: 4.8, SD: 0.9) and when N=20, average search length goes down to 38.4% (95% CI=[0.35,0.41], M: 7.6, SD: 2.28). These results indicate a trend to search less with increasing search length relative to the maximum time horizon. A test of mean group differences supports this conclusion, showing strong evidence for differences in relative search length between conditions for N=5 vs N=10: $^{SL}_{Diff}$ 0.12, HDI_{95} 0.11, 0.14 , BF_{10} 300 and for N=10 vs N=20: $^{SL}_{Diff}$ 0.1, HDI_{95} 0.08, 0.12 , BF_{10} 300 (R BayesFactor::ttestBF package, prior scale = medium; Morey et al., 2018). Participants performance (accumulated points per trial/ total amount, see Figure 3.9 B) increased significantly with extended time horizon, from an average performance of 82.8% (95% CI [0.82,0.83]) when N=5, to 85.7% (95% CI =[0.84,0.87]) when N=10 and 86.1% (95% CI =[0.85,0.87])) when N=20. This finding was supported by a t test producing Bayes Factors > 300 in favor of a difference in performance between time horizons.

Furthermore, we conducted a correlation analyses (R correlationBF package, prior scale = medium; Morey et al., 2018) to test if participants are stable in search length and performance across time horizons. The resulting Bayes factor > 300 indicates that there is extreme evidence for a correlation of search length across time horizons (between N=5 and N=10: $0.79\ 0.63, 0.85$ and between N=10 and N=20: $0.81\ 0.75, 0.88$). Furthermore, we found strong evidence that performance is correlated between time horizons (BF> 300, between N=5 and N=10: $0.63\ 0.59, 0.79$ and between N=10 and N=20: $0.82\ 0.76, 0.90$). These results suggest that participants have stable search tendencies across time horizons.

Given that participants choices are best captured by a model that assumes linear thresholds (we anticipate here the cognitive modeling results) we want to understand to which extent participants behave optimal under the constraints of linearity. Therefore, we calculated the best performing linear thresholds for each time horizon using grid search to find the parameter values ($_0$ and) that result in the highest reward. [2] Figure 3.9 (pink color) displays the relative search length by a decision maker who uses the best performing linear thresholds and reveals that relative search length decreases as well with increasing time horizons (N=5: 58% (95% CI=[0.57,0.59], M: 2.9, SD: 0.2) N=10: 53 % (95% CI=[0.52,0.52], M: 5.3, SD: 0.36), N=20: 48% (95% CI=[0.47,0.49], M: 9.6, SD: 0.6)). Compared to participants relative search length, a Bayesian t test indicates no evidence for a difference in search length when N=5 ($^{SL}_{Diff}$ 0.001, HDI_{95} $0.02, 0.02$, BF_{10} 0.17) but strong evidence

2 We used the same procedure as simulation study that describes the relationship between parameters and behavioral measures (Fig. 3.3, however we used a finer-grained grid to arrive at the best possible parameter value.

for a difference when N=10 ($^{SL}_{Diff}$ 0.08, HDI_{95} 0.06, 0.1 , BF_{10} 300) and N=20 ($^{SL}_{Diff}$ 0.11, HDI_{95} 0.09, 0.14 , BF_{10} 0.17). These results show that in longer sequences, participants search length is notably reduced compared to a policy that assumes best performing linear thresholds, by 8% when N=10 and by 11% when N=20. Performance of a best performing linear threshold model is increasing with extended time horizons (N=5: 87.2% (95% CI [0.86,0.88]), N=10: 90.3% (95% CI [0.89,0.91]), N=20: 92.7% (95% CI [0.92,0.93]). Participants thus achieve up to 92% of the performance when N=5, 91% when N=10 and 89% when N=20, which indicates of reduction of optimal adaptation in longer sequences. However, some subjects arrive at levels of performance competitive with the best performing linear threshold rule whereas others achieve only 47%. These findings suggest that there are individual differences on how people adapt to time horizon, while some participants may behave optimal relative to the best performing linear thresholds, others show sub–optimal adaptation to extended time horizons. The causes of individual differences in adaptive behavior may be found in the maladaption of cognitive parameters such as the initial aspiration level or its aspiration rate to time horizon, which will further investigate in the next sections.

Furthermore, Figure 3.9 displays that using the best performing linear thresholds (pink) reduces performance only slightly relative to the optimal model (yellow), by 1.1% when N=5, by 3% when N=10 and by 6.5% when N=20. This result reveals that an ideal adaptations of the linear thresholds to time horizon would lead to almost optimal performance and emphasizes the efficiency of the linear threshold strategy to solve the optimal stopping task.

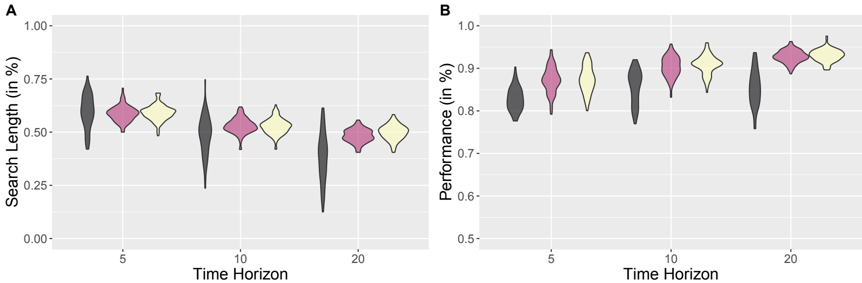

FIGURE 3.9: Violin plots of (A) Search length (in %: search length/total sequence length) and performance (in %: points/maximum points) for each time horizon (N=5, 10 and 20). Black color: Participant's data, pink color: Predictions from a model that uses best performing linear thresholds, yellow color: Predictions from a model that uses the optimal thresholds.

3.4.3 *Modeling results*

3.4.3.1 *Model comparison*

Using the same method as in Study 1, we applied the LTM, ITM and BOM to the participants' data to test their predictions against each other. The difference in the DIC between the LTM the BOM was 1356 in favour of LTM, indicating that the LTM is a better model to predict participants' choices in different time horizons. The comparison with the ITM revealed a lower DIC in favor of the LTM (a difference of 60). Figure 3.10 A shows the recovered thresholds for the LTM, the ITM and the BOM. Apparently, the LTM and the ITM thresholds are almost congruent when N=5 and 10. However, when N=20, the overlap of the thresholds is reduced, mainly due to a spike in the probability to accept higher ticket prices on position 19, which can not be captured by a linear threshold. While this spike seems in line with the BOM it is much less pronounced than predicted, thus indicating that the BOM, which

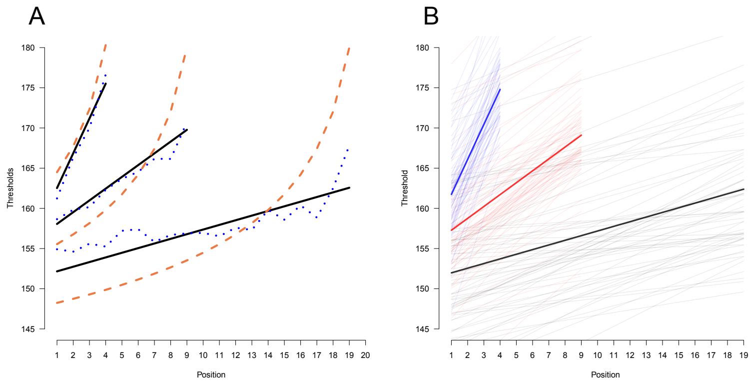

FIGURE 3.10: A: Estimated threshold parameters from the LTM (dark solid lines), the ITM (dotted blue lines) and the BOM (orange dashed lines). B: Individual-level threshold parameters for N = 5 (blue), 10 (red) and 20 (black). The three darker lines show the thresholds based on the group-level parameters.

	Group-mean parameter estimates				
LTM parameter	N=5	N= 10	N= 20	$N5$ $N10$	$N10$ $N20$
0	161.1 [159.4,162.7]	156.2 [154.2,158.2]	151 [148.6,153.3]	4.9 [3.5,6.2]	5.2 [2.1,8.3]
	4.4 [3.8,4.9]	1.48 [1.24,1.71]	0.58 [0.46,0.71]	2.9 [2.4, 3.3]	0.9 [0.6,1.1]
choice sensitivity	0.21 [0.18,0.23]	0.19 [0.16,0.21]	0.17 [0.15,0.19]	0.01 [-0.02, 0.04]	0.019 [-0.01,0.05]

TABLE 3.4: Posterior group-level mean parameters for each sequence length. 4*th* and 5*th* column: Difference between the posterior distribution (N5 - N10: 4*th* column, N10 - N20: 5*th* column). Note: Values in square brackets are the 95% credibility intervals (CI).

assumes biased optimal threshold parameters, is not capable to capture participants' choices. Thus, the general pattern is in line with the LTM assumptions, leaving only the question open, whether the final position, with fixed time horizons, reflects a special case revealing itself with longer time horizons, which we seek to address in future studies.

3.4.3.2 *Best-fit parameters*

When assuming that the three LTM parameters have psychometric properties, the comparison of the best-fit parameters between the time horizon conditions can shed light on the processes involved in adaptive behavior. The estimated by-position thresholds for each time horizon condition are plotted in Figure 3.10 B, along with the thresholds obtained with the individual-level parameters. Table 3.4 summarizes the group mean parameter values of $_0$, and in each time horizon condition, and shows that both $_0$ and are reliably reduced with increasing time horizon. Column 4 and 5 report the mean of the difference of the population parameter's posterior distributions and indicates strong evidence for adaptation of the first aspiration level ($_0$) and its adjustment across time ( ) across time horizon (difference is credibly different from zero). However, there is no evidence that choice sensitivity ( ) is modified between conditions (0 lies within the 95% confidence interval). This result supports the hypothesis that both $_0$ and capture the participants' adaptation to the total number of options in the sequence. Furthermore, it confirms the LTM's predictions (simulation study, Fig. 3.3) indicating that in order to maintain optimality, both the first thresholds and its adjustment over position should be reduced in longer time horizons.

To understand if there is a systematic way how people adapt their thresholds to time horizon, we looked at the quantitative properties of the threshold adaption across time horizons. We find that doubling the sequence size from N=5 to N=10 leads to a reduction of the initial threshold by a mean value of 4.9 [3.5,6.2]. Likewise, when doubling the sequence size again from N=10 to N=20, the initial threshold is decreased by a mean value of 5.2 [2.1,8.3] (see column 4 in Table 3.4). The comparison of the initial threshold' adaption between time horizon conditions (${}_{0}^{N5}$ ^{N10} - ${}_{0}^{N10}$ ^{N20}: M= -0.02, CI=[-0.21 0.18]) indicates no difference (0 lies within the 95% HDI). Furthermore, we investigated how participants adapt between time horizons. The rate is reduced with increasing sequence length, leading to a very little adjustment of the initial threshold when N=20. Our explorative analysis suggests that is updated according to the time horizon such that doubling the sequence size reduces the adjustment to one third of the original (and vice versa: increased by 3 when sequence length is halved). Although it appears that participant's initial thresholds and adjustment rates are adapted systematically to time horizons, further research is needed to investigate which cognitive functions could account for this non–linear adaptation.

To understand participants' decrease of search length and performance in extended time horizon relative to the best performing linear threshold model, we compared the participants' thresholds with the best performing linear thresholds (see Supporting Material, Fig. 3.14). The best performing linear threshold rule predicts an adaptation of the initial thresholds proportional to time (doubling the sequence leads to a decrease of the initial threshold by a constant amount) and the adjustment across position (doubling the sequences leads to halving the adjustment rate). Therefore, although participants seem to follow the same regularities in adapting the initial threshold and its adjustment rate to time, they adapt too little to extending time horizons. Accordingly, whereas participants' decision thresholds are quite similar to the best performing linear thresholds when N=5, they deviate more strongly when N=10, and even more when N=20.

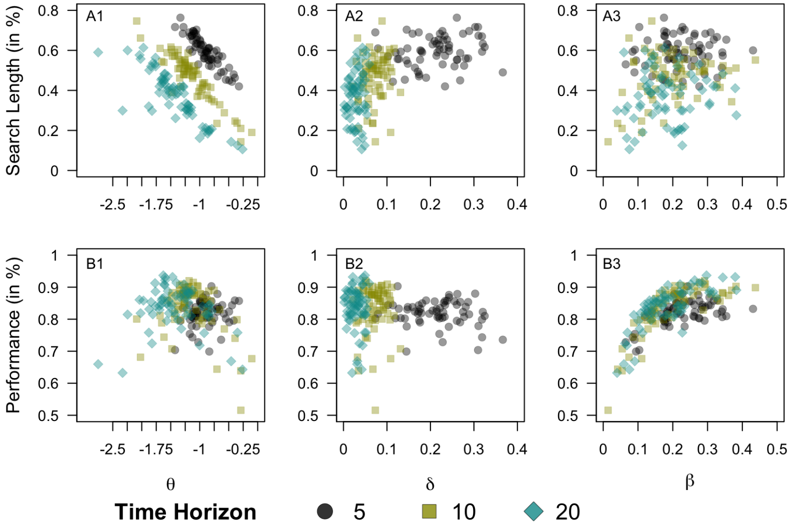

FIGURE 3.11: Study 2: A1-A3: Scatterplots of search length (in %, search length/total sequence length) vs parameters ($_0$, and). B1-B3: Performance (in %, points/maximum points) vs parameters. Black: N=5, light green: N=10, dark green: N=20.

3.4.3.3 *Capturing behavioral patterns*

To find out which of the estimated LTM parameters predict search length or performance, we performed a correlation analysis. The scatterplot in Fig. 3.11 shows the results for each condition, as indicated by the symbol colors. Fig. 3.11 A1 suggests a negative linear relationship between the first thresholds $_0$ and search length in all time horizons, and also indicates the systematic differences on both search length and threshold estimates between conditions. A Bayesian linear regression analysis (R BayesFactor::lmBF package, prior scale = medium; Morey et al., 2018) supports the evidence of these assumptions: for each time horizon condition, that is N=5, 10 and 20, $_0$ is an important predictor of search length and accounts for 76%, 81% and

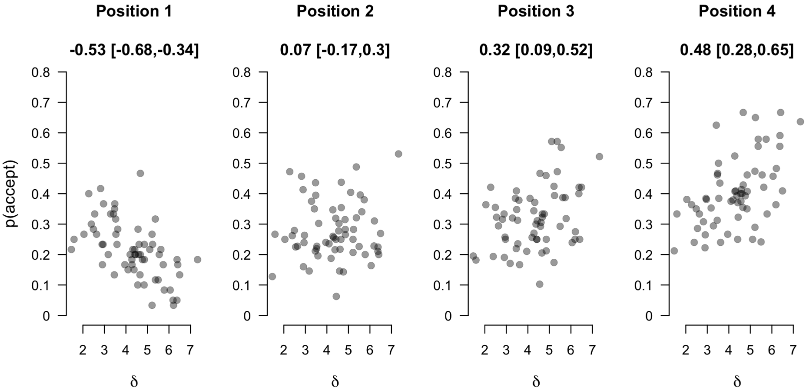

FIGURE 3.12: Scatterplots of the relationship between and the probability to accept on each position (N=5, see Supporting Material Fig. 3.15 for N=10 and Fig. 3.16 for N=20). Cor: Correlation coefficients [HDI: 95% High Density Interval]

66% (respectively) of search length' variation (see Supporting Material Tab. 3.10 and 3.11 when N=5, Tab. 3.14 and 3.15 when N=10, and Tab. 3.18 and 3.19 when N=20). These correlation results are in line with those from Study 1 and suggest that the initial aspiration level is the main predictor for participant's search length in different time horizons. Furthermore, Fig. 3.11 A2 displays a non-linear interaction effect between the manipulation of time horizon and , whereas the time horizon groups (N=5, 10 or 20) can be either differentiated based on the values or search length (for search length rather indicated as a variance reduction).

The scatterplot displaying the relationship between performance and parameters (Fig. 3.11 B1) suggests an inverse U-shaped relationship between $_0$ and performance, consistent with the prediction from the simulation study of the LTM (Fig.

3.3). Furthermore, exhibits a linear relationship with performance (Fig. 3.11 B2). A regression analysis reveals that in short sequences (N=5), the best model predicting performance includes only the parameter (explaining up to 28% of its variation, see Supporting Material Table 7 and 8), whereas in longer sequences (N=10 and 20), the best model is comprised of both and $^{2}_{0}$, accounting for about 45% and 44% (respectively) of the variance in performance (see Supporting Material, Tab. 3.17 and 3.18 when N=10 and Tab. 3.20 and 3.21 when N=20). This result indicates that with increasing time horizon, the initial threshold becomes predictive of performance, whereas setting it too high or too low leads to poorer performance. Additionally, (choice sensitivity) is a further strong predictor of the variation in performance, where higher values in result in higher payoff. For additional comparisons between high versus low performing individuals see Supporting Material, Figure A1 Fig. 3.13

While the overall impact of in the above analyses rather concern differences in time horizon, it further reveals individual choice dynamics in relation to the probability to accept ticket prices over sequence positions. Figure 3.12 shows corresponding scatterplots for the time horizon of N=5 (one for each sequence position; plots for N=10 and N=20 show similar trends, see Fig. 3.15 (N=10) & Fig. 3.16 (N=20)). As can be seen, on position 1, is negatively correlated with the acceptance probability, which becomes increasingly positive with ongoing positions. This result illustrates individual differences in adaptive behavior over positions by the functional characteristics of the LTM. That is, weaker adaptation rates imply higher initial thresholds and, thus, higher acceptance probabilities in the beginning (i.e., an already high threshold leads to early accept choices by definition, without requirement of strong adaptation). Consequently, the change in this relation with ongoing positions indicates individual differences between participants who accept early, and those who accept late. The latter participants, as can be seen, seem to drive the reversal of the correlation, as they become more and more likely to accept ticket prices (regardless of their values). Thus, the adaptation rate could be an interesting measure to further investigate individual differences in the cognitive functions underlying behavioral adaptation in future studies.

3.4.3.4 *Stability of LTM parameters*

A further goal of this study is to asses the reliability of the search processes, captured in the LTM parameters, across tasks. A correlation analysis (R BayesFactor::correlationBF package, prior scale = medium; Morey et al., 2018) of the parameters between time horizon conditions (see Table 3.5) reveals positive correlations between the same parameters. $_0$, reflecting participants' aspiration level before search, reveals the strongest correlation (between 0.61-0.8), pointing to a stable mechanisms across conditions. The correlation coefficient of is smaller (between 0.27-0.51) across conditions, which could be an indication that this parameter is estimated with higher uncertainty. Note that – as found in Study 1 – the first threshold ($_0$) and the adjustment rate across search ( ) are negatively correlated to some extend, showing a typical parameter dependency as in all regression-like (intercept plus slope) models.

	5 0	10 0	20 0	5	10	20	5	10
5 0	1							
10 0	**0.80 [0.69,0.87]**	1						
20 0	**0.61[0.45,0.74]**	**0.73[0.61,0.83]**	1					
5	-0.25[-0.4,-0.2]	-0.12[-0.3,-0.1]	-0.16[-0.4,0.06]	1				
10	-0.2[-0.4,0.02]	-0.14[-0.34,0.07]	-0.19[-0.4,0.02]	**0.51[0.32,0.67]**	1			
20	-0.2[-0.41,0.01]	-0.11[-0.32,0.11]	-0.15[-0.36,0.07]	**0.31[0.1,0.5]**	**0.27[0.04,0.47]**	1		
5	0.00[-0.2,0.2]	-0.15[0.36,0.07]	-0.08[-0.3,0.14]	-0.02[-0.4,-0.02]	-0.12[-0.3,0.09]	0.00[-0.23,0.22]	1	
10	0.00[-0.2,0.2]	-0.05[-0.27,0.17]	-0.05[-0.28,0.17]	-0.13[-0.35,-0.1]	0.00[-0.22,0.2]	-0.05[-0.27,0.18]	**0.62[0.45,0.74]**	1
20	0.00[-0.2,0.2]	-0.05[-0.2,0.1]	0.07[-0.15,0.3]	-0.06[-0.28,0.16]	-0.1[-0.3,0.12]	-0.05[-0.28,0.16]	**0.66[0.51,0.77]**	**0.74[0.61,0.83]**

TABLE 3.5: Correlations between LTM parameters when N=5, 10 and 20, with 95% HDI (bold numbers: BF > 150)

3.4.4 *Discussion Study 2*

The main purpose of the second study was to investigate the systematic influence of time horizon on adaptive choice behavior in optimal stopping tasks (searching the cheapest ticket) and their cognitive description using the LTM (Baumann et al., 2020). We found, participants stopped their search earlier in longer sequences (relative to the total sequence length) and improved their performance, suggesting that they were sensitive to the time horizon. Moreover, compared to shorter time horizons, accept probabilities were lower in the beginning of search and increased less strongly across positions. When modeling the participants' choices with the LTM using a Bayesian hierarchical approach, we showed that the effect of time horizon was captured by changes in initial decision thresholds (aspiration level), and the adaptation rate of the thresholds with ongoing sequence positions (i.e., with growing number of rejections). More specifically, with longer time horizons, the initial aspiration level increased reflecting more frequent initial rejections in longer time horizons. Similarly, the adaptation rate of the threshold across sequence positions was weaker for longer time horizons reflecting a slower change in the probability of accepting ticket prices.

However, a comparison with a best performing linear threshold model (using grid search to find best performing linear thresholds) reveals that not all participants adapted their threshold perfectly and thus reach sub–optimal performances, yielding corresponding average trends in all three time horizon conditions. Especially in longer sequences, as for N=10 and 20, participants tended to set their initial aspiration level too high and thus to terminate their search too early. Importantly, we found that this pattern is more pronounced with a growing sequence length (N=20).

We further examined the relation between individual differences in the participants' parameter estimates to search length and performance. In all three conditions, we found that individual differences in search length were almost exclusively related to the initial threshold (aspiration level prior to search). This was also functionally related to the adaptation rate, such that strong adaptation of initially low thresholds predicted an increase of the probability to accept tickets at later positions. Furthermore, the first threshold is also predictive of performance, showing an inverse U-shaped relationship, with extreme (too high or too low) values leading to lower payoff (when N=10 and 20). These findings correspond to the results of Study 1 and also to results from previous literature (Lee et al., 2004; von Helversen et al., 2012; Zwick et al., 2003). Importantly, it suggests that the primary factor affecting search length and performance is largely determined by the initial threshold reflecting participants' aspiration level prior to search.

In addition, a strong predictor of performance was the LTM parameter of choice sensitivity, representing the participants' responsiveness to the distance between the aspiration level and the actual price. A higher choice sensitivity reflects more deterministic use of the decision threshold, while lower sensitivity reflects probabilistic responding. Interestingly, the analysis of search behavior in related domains has revealed that sensitivity parameters in the BART (Lejuez et al., 2002) or the Iowa gambling task (Bechara et al., 1994) are related to real world risky behavior (Guan et al., 2020; Pleskac, 2008; Stout et al., 2004; Wallsten et al., 2005). Consequently, one might predict that sensitivity in an optimal stopping task might as well be able to identify real-world risk takers which merits further investigation in future studies.

We find moderate to strong consistency between the same parameters across time horizon conditions which provides evidence for stable individual cognitive processes beyond the different time scales. However, the correlation of the initial thresholds between tasks is stronger than the threshold's adjustments over time. The individual choice sensitivity, also referred to as "decision noise" is strongly associated between conditions, showing that participants are consistently sensitive to the distance' evaluation between the actual price and their threshold across time horizon conditions.

3.5 GENERAL DISCUSSION

The goal of our study was to learn about the aspects of adaptive behavior to outcome variance (Study 1) and time horizon (Study 2) in optimal stopping tasks. Our investigations were based on the idea, that decision makers can be described by three parameters, the initial aspiration level, its adaptation over time, and choice sensitivity (response determinism). To measure these parameters, we used the linear threshold model (LTM, (Baumann et al., 2020)) which assumes that decision thresholds are linearly adjusted over time.

In Study 1, we found that people show an almost perfect adaptation of their decision thresholds to the sampling distribution. This finding is in line with the normalization hypothesis which assumes that the value of an option is equal to its rank within the given sample. In Study 2, we found that people adapt their choices to the time horizon by decreasing both the first aspiration level and its adaptation rate across positions with longer time horizons.

In the following we discuss the implications from perspectives of mental representations (normalization of values) and other theoretical approaches to adaptive behavior in sequential search.

3.5.1 *Adaptivity*

The first study shows that outcome variance has an effect on peoples decision thresholds, with increased thresholds in lower variance environments. However, the decision thresholds are practically identical on a normalized scale (0, 1), showing that people accept prices within the same percentile on each position, regardless of the variance. This result suggests that the price' value is assessed on the its percentile rank within the sample rather than on its absolute number. Our finding is thus in line with the normalization hypothesis, which assumes that the value of an option is computed under a normalized code and corresponds to its relative position in the distribution of options (Rangel et al., 2012). It is also consistent with the idea that the value of a single option is calculated by comparing it to a sample of attribute values drawn from memory and is its rank within the sample (Stewart et al., 2006).

Consequently, the finding that the individual acceptance level is formulated on a percentile level (e.g. accept all prices on the first position that belong to the 12% lowest prices within the sampling distribution) and is consistently adapted to the sampling distribution has implications for optimal choice behavior in different search conditions. In particular, people are expected to use the same percentile threshold in other distributional environments, for example when values are sampled from a left– or right–skewed distribution. Indeed, some studies (Baumann et al., 2020; Guan et al., 2018) have found evidence that that people used higher thresholds in an environment with many good alternatives compared to an environment with only few good options, which corresponds to the qualitative prediction of the normalization hypothesis. Additionally, in accordance with our findings, Lee et al. (Guan et al., 2018) have shown that thresholds which people use when searching for the maximum (when values $\quad beta\ \ 4,2\ $) can be modeled in the same way as the thresholds that people use when searching for the minimum (values $\quad beta\ \ 2,4\ $). Therefore, future work should attempt to replicate our analysis in setting where options are sampled from different distributions and thus investigate if the normalization hypothesis represents a stable cognitive mechanism of the valuation process in sequential search tasks.

A further questions arising under the normalisation hypothesis relates to the valuation of options when the sampling distribution is not (or partly) known. Indeed, there is evidence that search behavior in non–stationary environments are highly dependent on the history of samples during search (e.g. Brickman, 1972; Corbin et al., 1975). These studies differ from our tasks regarding the knowledge about the sampling distribution: Whereas in our study, the sampling distributing was intensively learned before search, in these studies it was experienced during search. In this case, the valuation of options happens relative to the options seen so far and thus can lead to systematic biases in the threshold calculation. For example, if options at the beginning of the sequence belong to the lower end of the distribution and thus are not representative of the full sampling distribution, calculating a threshold based on percentiles ("I am willing to accept the 10% best options") would lead to decision thresholds that are too low relative to the optimal thresholds. In order to understand the impact of small samples on the valuation of options in optimal stopping tasks, one would need to extend cognitive models of optimal stopping search by incorporating a learning component, what we believe would be a useful direction for future research.

The second study shows that time horizon affects the decision thresholds, whereby extended time results in lower first thresholds and lower adjustment rates across positions. In other words, participants are less selective at the beginning of the the sequence in shorter time horizons and also adapt their acceptance rate more strongly across positions. This strategy is intuitively correct because as the as the task gets shorter, there are less chances along the way for a good option to appear. The qualitative adaption of participants' initial thresholds and its adjustment over time is identical to the predictions of the LTM simulation about optimal parameter adaption across time. However, with increasing time horizons, participants fail to adapt the parameters optimally, leading to reduced search length and performance.

Our investigation further suggests that participant's initial thresholds and adjustment rates are adapted systematically to time horizons in a non–linear way, indicating that they might follow scaling regularities across time scales. Similarly, a study using a simplified explore–exploit task (Sang et al., 2020) has shown that participants' linear decreasing thresholds are declined in proportion to the length of the game. Indeed, the best performing linear thresholds predicts a proportional adaptation of the initial threshold and its adjustment rate and in this sense, participants' adaptation is optimal. Nevertheless, participants adapt too little in extended time horizons, leading to stronger deviation from optimality. Further cognitive mechanisms, which were not considered in our analysis, such as risk preference or search costs, could be reasons for the sub–optimal adaption to extended time scales, which we elaborate in the following section.

Some studies have suggested that people use biased decision thresholds and thus search too little due to humans risk averse preference (Bhatia et al., 2021; Schotter et al., 1981; Schunk, 2009). Unlike the risk neutral best performing linear threshold model, human decision makers are often found to be risk averse (Pedroni et al., 2017, e.g.), thus preferring the safe over the risky option. Based on this assumption, suboptimal adaptation of higher initial thresholds to longer time horizons may reflect risk aversion that may become more visible with increased time horizons. However, findings regarding the impact of risk aversion on optimal stopping behavior are ambivalent. One study reports that there is no relationship between search length in an optimal stopping task and risk preferences elicited in a series of lottery tasks (Schunk, 2009). On the other side, a recent study about optimal sequential search with recall found evidence that in this paradigm, risk aversion is one of the main factors guiding people's behavior (Bhatia et al., 2021). Moreover, Sonsino et al.

(Sonsino et al., 2002) suggested that the probability of choosing a given option decreases with the relative complexity of that alternative. In that sense, participants might perceive a higher time horizon as more risky or more complex and thus devalue the option of search more heavily, leading to behavior that deviates from optimality.

Alternatively, Seale and Rapoport (Seale et al., 2000) hypothesized that stopping too early in the optimal stopping task could result from the existence of endogenous search costs, i.e. psychological costs arising from considering and evaluating options. Although our experiment used a somewhat different design, participants may have been motivated to end the experiment as quickly as possible and thus considered the time spent in observing and reviewing the prices as a kind of search cost. This idea is further supported by the literature on the accuracy-effort trade off reported by Payne et al. (1993) and could also explain why search was more strongly reduced with longer time horizons. In our experiment, search costs may be integrated into the evaluation of the decision thresholds, while more alternatives sequences lead to higher search costs (more steps to go through all the alternatives) and thus to a higher impact on decision thresholds.

An simple explanation for the week adjustment of the threshold across position (relative to the adjustment of the best performing linear thresholds) in extended time horizons might be that they result from their originally higher initial thresholds. However, the origins of such a maladaption of the thresholds within a sequence might be that participants misrepresent the underlying outcome distribution. In a sequential search paradigm such as the optimal stopping task, the outcome distribution changes dynamically and thus has to be recalculated on each time step. Since people seem to lack the ability to infer many characteristics of aggregated outcome distributions correctly (Benartzi et al., 1999; Klos et al., 2005), simpler strategies are used solve the task (see also Speekenbrink et al., 2015; Wilson et al., 2014, for similar discussion for bandit tasks). Indeed, results from the BART (a non–stationary environment) showed that the best fitting model to two datasets assumes a stationary representation (Bishara et al., 2009; Wallsten et al., 2005), showing that people simplify the balloons' exploding behavior. Preliminary results from simulation studies of optimal stopping search indicate that subjects might believe that the environment in an sequential search tasks corresponds to a hypergeometric distribution (sampling options without replacement) which would lead to the observed choice patterns across time scales. Future work is needed to incorporate such mental representations and updating mechanism into optimal stopping models to shed light on the underlying mechanisms that drive the threshold adjustment across search.

3.5.2 *Individual differences in choice behavior*

Most research on optimal stopping search has focused on data which reflect only the end product of the decision process, such as search length and accuracy. In contrast, the present study provides valuable insights into the individual differences in cognitive processing which led a subject to exhibit a particular choice.

Our result indicate that the initial aspiration level, set prior to search, is the main factor predicting participants' search length. Indicated by the negative linear correlation coefficient across the tasks in both studies, higher initial aspiration levels lead to less search and vice versa. Importantly, the rate of the threshold's adjustment during search has no additional impact on search length, indicting that participants' modification of their aspiration level during search has little impact on how long their search. Furthermore, this result suggests that a too high aspiration level prior to search may serve as an explanation for the general finding of undersearch (Hey, 1982; Rapoport et al., 1970; Schotter et al., 1981) and thus deserves special consideration.

The relationship between the initial threshold and performance in non–monotonic, indicating that too extreme values (too high or too low) in the initial threshold lead to worse performance, and thus reveals that not only too little, but also too much search can be sub–optimal. However, despite the finding that participants perform surprisingly good throughout all the tasks, we find that a longer sequence length (20 alternatives, Study 2) lead to more pronounced undersearch for a larger part of participants (elongated longer tail in search length distribution) and thus lead to worse performance. However, some participants that achieve almost optimal performance across the tasks (Study 2) and thus adapt their decision thresholds perfectly to the time horizons. Taken together, this result shows that participants' not only differ in their aspiration levels but also in how well they adapt them across time horizons. What could be the reason for these individual differences ? A possible explanation would be that the tendency to delay future outcome could play a role in how people adapt to time horizon. However, Meyers et al. (Meyers et al., 2020) has recently shown that delay discounting is inconsistently predictive of search behavior in a secretary problem. Other studies have identified some individual characteristics that could account for individual differences in adaptive behavior to time horizons, such as working memory capacity (people with higher capacity tend to explore more) (Hills et al., 2012), age (older people may adapt less) (Mata et al., 2013; Rydzewska et al., 2018)) or people with depression may explore more (von Helversen et al., 2011). An interesting problem for future research is whether the observed individual differences in adaptation are related to other measures such as delay discounting or working memory.

Participants' response sensitivity is highly predictive for their performance, with higher values (and thus more determinism) leading to better performance. Indeed, high performing participants' choice curves are more consistent compared to low performing participants thus showing less deviance from the intended choice (see Fig. 3.13). However, we find no relationship between participants' choice sensitivity and their search length. This result could shed light on the complex relationship between risk preferences and search behavior. On one side, studies relating risk preferences elicited in lottery tasks with search length in optimal stopping tasks have not found any correlation (Schunk, 2009). On the other side, results from related sequential risk-taking tasks, such as the the Angling Risk Tasks or the Balloon Analogue Risk Task report that sensitivity to the outcome evaluation as being predictive for individual differences of real-world risk takers (Pleskac, 2008; Wallsten et al., 2005). Exploring the relationship between choice sensitivity in search tasks and risk preferences would be a useful direction for future research.

3.5.3 *Stability of parameters*

Finally, we measured consistencies in the LTM parameters, reflecting the initial aspiration level (captured in the initial threshold), its adjustment rate and the choice sensitivity, across contexts. We found that the same parameters are stable across changing contexts such as different variances and time horizons and suggests that these parameters thus reflect reliable processes that are involved when solving optimal stopping tasks. Nevertheless, the correlations are far from perfect. Besides genuine instability, these may reflect variance in how people respond to the changes in the task, i.e. the variance or the time horizon. Additionally, measurements of the parameters become noisier with longer sequences leading to lower correlation values when N=20. Research on exploration-exploitation trade-offs has frequently emphasized that individual search might underlie a general mechanism that affects search across different domains and tasks (Hills et al., 2008; Hills et al., 2015; Mata et al., 2015; Pirolli, 2007). However, attempts to link behavior across different exploration-exploitation tasks have been unsuccessful (e.g. von Helversen et al., 2018). Stable cognitive process parameters in the sequential search tasks may thus help to discover the common mechanism of search across different exploration-exploitation tasks.

3.6 CONCLUSION

We inquired into whether and how context task variables affect optimal choice behavior by identifying and quantifying the cognitive processes involved in their adaption strategies. Results show that people adapt their decision thresholds to variance and time horizon. The adaption to variance leads to identical thresholds on a percentile scale, suggesting that the value of an option is determined by the rank within the pool of alternatives. Further, people adapt sub–optimally to time horizons, leading to pronounced undersearch in extended time horizons. This work contributes to the understanding of the adaptive processes that underlie sequential decisions and thus will help quantify the conditions under which people may succeed or fail in such tasks.

3.7 SUPPORTING MATERIAL: ADAPTIVE BEHAVIOR IN OPTIMAL SEQUENTIAL SEARCH

Text

1. Calculation of optimal thresholds

We describe the calculation of optimal thresholds applied to our scenario, where payoff is proportional to the chosen value and the goal is to find cheapest ticket price. We first derive the optimal solution mathematically based on the paper of Gilbert and Mosteller (Gilbert et al., 1966, Section 5b) and further provide a more intuitive explanation.

Let's assume a sequence of n ticket prices that are drawn from a standard normal distribution with density:

$$f \;\; x \qquad \frac{1}{\overline{2}}\,e^{\;\frac{1}{2}\,t\,\overline{}^{\,2}} \tag{3.5}$$

and the goal is to find the lowest ticket price in this sequence.

The *optimum strategy* for this task is:

If n 1, the decision maker is forced to accept the ticket. The threshold on the last position (T_1) is set to and ticket prices below this threshold are accepted:

$$T_1 \tag{3.6}$$

Therefore the expected price of the last ticket (P_1) is the mean ( ) of the distribution.

If n 2, the decision maker decides to keep the first option or to reject it and to go on to the second one. If he goes on, his expected ticket price is . Therefore he keeps the current one, x, if x , rejects it if x and is indifferent if x . Therefore, the expected price of the last ticket (P_1) is also the threshold for the second last option:

$$T_2 \quad P_1 \tag{3.7}$$

Than for n 2, his expected price (P_2) is:

$$P_2 \qquad {}^{P_1} f\ x \quad xdx \quad P_1 \quad {}_{P_1} f\ x\ dx \tag{3.8}$$

The remaining terms of the sequence can be computed in a recursive manner. For each n, the decision maker accepts the ticket if it is lower than the expected price of the remaining n 1 tickets (x $P_{n\ \ 1}$) but rejects if the ticket is higher than the remaining expected price ($x \quad P_{n\ \ 1}$) therefore the threshold on the n-th position (T_n) is:

$$T_n \quad P_{n\ \ 1} \tag{3.9}$$

Accordingly the expected price (P_n) is :

$$P_n \qquad {}^{P_{n\ \ 1}} f\ x \quad xdx \quad P_{n\ \ 1} \qquad {}_{P_{n\ \ 1}} f\ x\ dx \tag{3.10}$$

Intuitive explanation

The optimal thresholds T_n for maximising the payoff is calculated working backward from the last ticket price: The threshold of the final item (T_1) is , because the rules of the task stipulate that the final item must be accepted if no earlier item has been chosen. The thresholds for the previous items are determined by working backward from the final item, using conditional expectations. First, we calculate the expected value of the final item (P_1). For the last item, this is the expectation of the overall probability distribution from which the options are sampled. Therefore, to maximize expected reward on the second last position, one's policy should be to accept a particular option if it is better (in our case smaller) than the expected reward if one continues under the optimal policy. The second-to-last item should be accepted if its value is smaller than the expected value of the final item. This means that the threshold of the second-to-last item (T_2) is the expected value of the last item (P_1).

The expected value of the second-to-last item (P_2) is the expected value of the part of the probability distribution that is better (in our case smaller) than the threshold (T_2) for the second-to-last item. The probability of this expected value is the area under the probability distribution that is better than this threshold. The overall expected reward at the second-to-last position (P_2) (and therefore the threshold for the third-to-last item (T_3)) is calculated as follows: we multiply the expected value for the second-to-last item with its probability plus the expected value of the last item multiplied with its probability (which is equal to 1 minus the probability of the second-to-last item). The remaining thresholds are calculated in the same way.

2. Biased Optimal Model

The *Biased Optimal Model (BOM)* is based on the Bias-from-Optimal threshold model proposed by Guan et al. (2015), assuming that humans are using thresholds that deviate systematically from the optimal thresholds.. The optimal thresholds t_i for each position i are derived by determining the expected reward of the remaining options (derivation in (Gilbert et al., 1966) and in Appendix Text 1). The model entails a systematic bias parameter that reflects the divergence of the human threshold from the optimal one. Additionally, the thresholds depend on a parameter that determines how much their bias increases or decreases as the sequence progresses.

$$ {}_{i\;1} \qquad {}_{i} \qquad i, \tag{3.11} $$

When and are set to 0, the thresholds represent the optimal thresholds that lead to best performance. This model is therefore defined by three free parameters, , and the choice sensitivity .

Methods

- *Linear Threshold Model (LTM)*:

$_j \quad N \quad _g, \quad _g$

$_j \quad N \quad _g, \quad _g$

$t_{0j} \quad N \quad ^{t_0}_g, \quad ^{t_0}_g$

$_g \quad N \; 0, 100 \; , \quad _g \quad U \; 0.1, 10$

$_g \quad N \; 0, 100 \; , \quad _g \quad U \; 0.1, 10$

$^{t_0}_g \quad U \; 100, 200 \; , \quad ^{t_0}_g \quad U \; 0.1, 10$

- *Biased Optimal Model (BOM)*:

$_j \quad N \quad _g, \quad _g$

$_j \quad N \quad _g, \quad _g$

$N \quad _g, \quad _g$

$_g \quad N \; 0, 100 \; , \quad _g \quad U \; 0.1, 10$

$_g \quad N \; 0, 100 \; , \quad _g \quad U \; 0.1, 10$

$_g \quad N \; 0, 100 \; , \quad _g \quad U \; 0.1, 10$

- *Independent Threshold Model (ITM)*:

$_j \quad N \quad _g, \quad _g$

$$t\ p_{\ j} \quad N \quad {}_{g}^{t\,p}\ ,\ {}_{g}^{t\,p}$$

$$_{g} \quad N\ 0,100\ ,\ \ _{g} \quad U\ 0.1,10$$

$$_{g}^{t\,p} \quad U\ 0,220\ ,\ \ _{g} \quad U\ 0.1,20$$

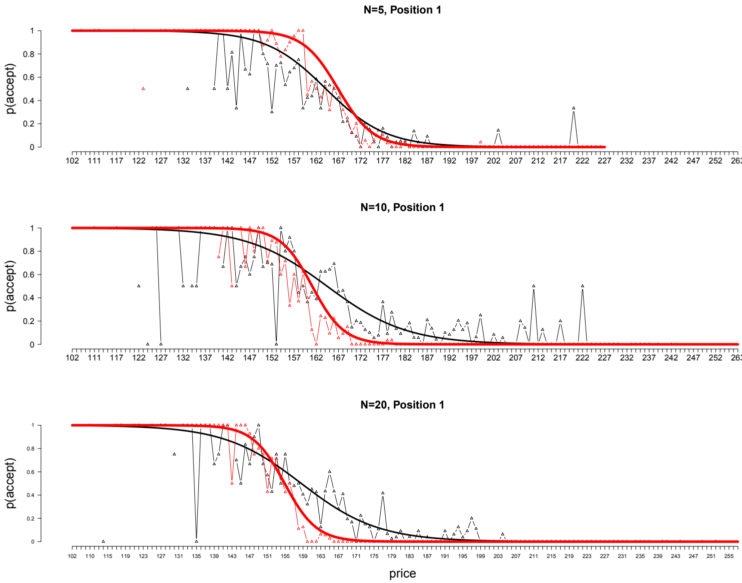

FIGURE 3.13: Choice curves split into participant that belong to the 20% highest (red lines) vs the 20% lowest performing group (black lines), for each search length and on the first position. Well performing participants exhibit higher determinism in choices than low performing participants. Solid lines are logistic curves fitted to the data (red line: fitted to data of participants in high performing group, black line: fitted to data of participants in high performing group.)

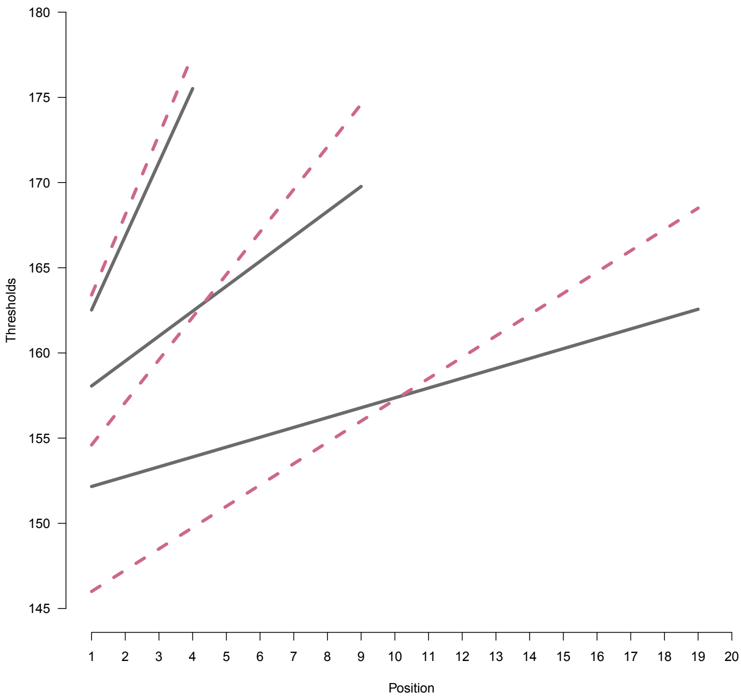

FIGURE 3.14: Thresholds for each time horizon (N=5, 10, 20). Black solid lines: Estimated threshold parameters from data (LTM). purple dotted lines: best performing linear thresholds (using grid search to find the best performing parameter values ($_0$ and))

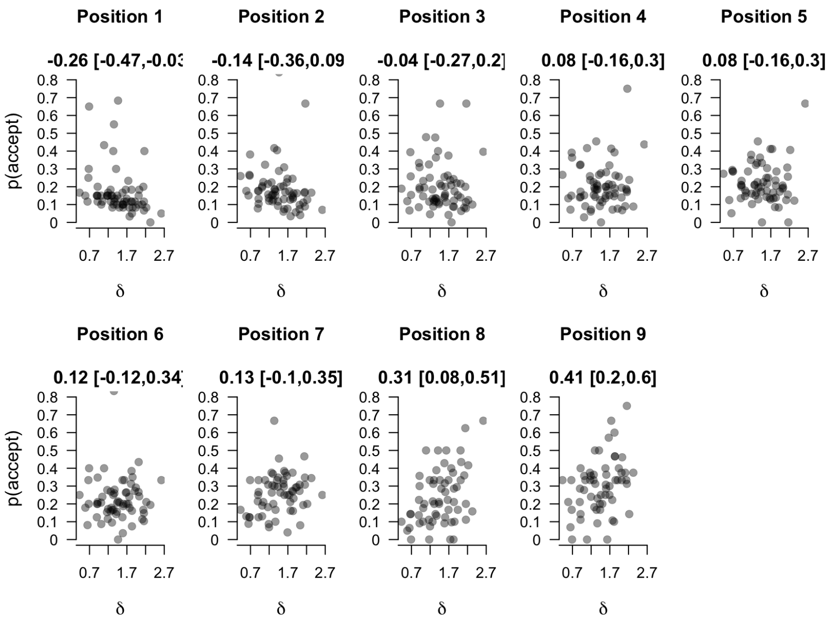

FIGURE 3.15: Study 2: Time horizon =10: Relationship between probability to accept and parameter across position. Scatterplots and correlation coefficients, HDI: 95% high density interval.

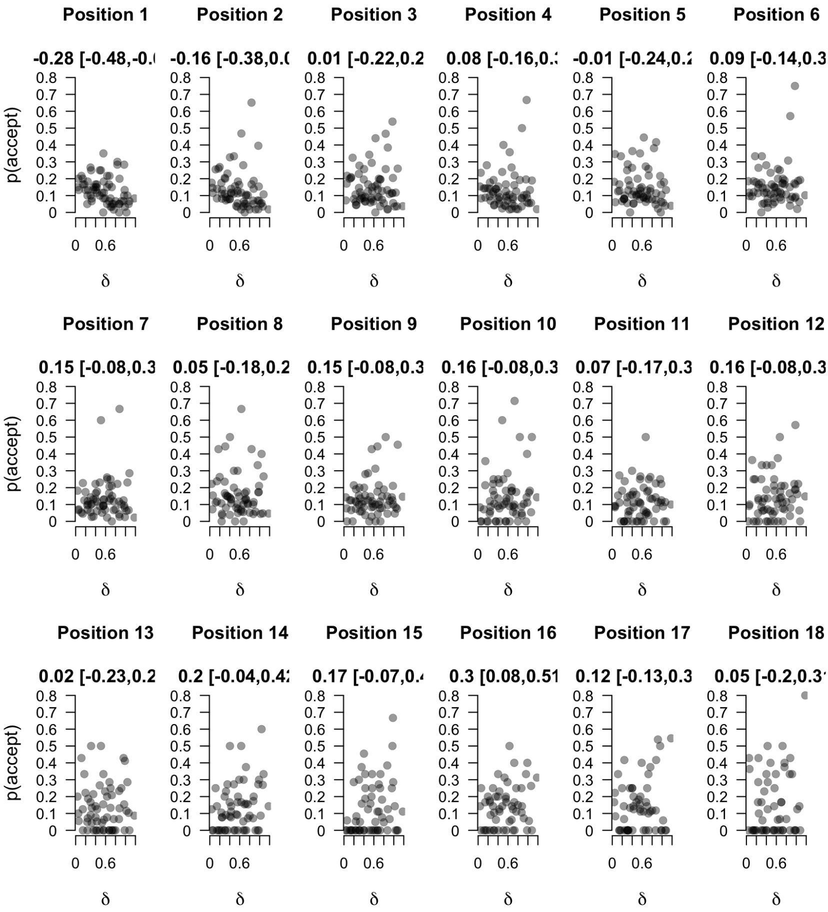

FIGURE 3.16: Study 2: Time horizon =20: Relationship between probability to accept and parameter across position. Scatterplots and correlation coefficients, HDI: 95% high density interval.

Models	B_{m0}	B_{mf}
+ +	$10x10^{181}$	1
+ + + condition	$3x10^{180}$	0.3
+	$2x10^{168}$	$2x10^{14}$
+ + condition	$5x10^{16}$	$5x10^{16}$
+	$5x10^{40}$	$5x10^{20}$

TABLE 3.6: Study 1: Ranking of the different models for the analysis of the search length using Bayes Factors. Selected model is in bold. B_{m0}: Bayes Factor between the model and the null model (null model: intercept only). Bmf: Bayes factor of each model relative to the full model (full model: SL + + + condition). Only the five models with the highest support are presented.

Parameters			R^2
intercept	$0.8x10^4$		
	-0.98 [-1.0,-0.95]	0.89 [0.91,0.87]	0.80
	-0.16 [-0.19,-0.12]	0.21 [0.11,0.29]	0.04
	0.26 [0.23,0.30]	0.11 [0.16,0.20]	0.01

TABLE 3.7: Study 1: Coefficient estimates ( ) for each of the predictors in the best fitting model in Table 1, along with the 95% HDI. : Correlation (plus 95% HDI) with search length. R^2: Proportion of variance in search length that is predictable from the independent variable

Models	B_{m0}	B_{mf}
^{2} + +	$4x10^{74}$	1
^{2} + + + condition	$1.5x10^{74}$	0.4
+ ^{2} + +	$3x10^{73}$	0.07
+ ^{2} + + + condition	$1.2x10^{73}$	0.03
+	$3x10^{70}$	0.0005

TABLE 3.8: Study 1: Ranking of the different models for the analysis of the performance using Bayes Factors. Selected model is in bold. B_{m0}: Bayes Factor between the model and the null model (null model: intercept only). Bmf: Bayes factor of each model relative to the full model (full model: SL + + + condition). Only the five models with the highest support are presented.

Parameters			R^2
intercept	$0.4x10^{4}$		
^{2}	-0.38 [-0.43,-0.32]	-0.65 [-0.59,-0.73]	0.42
	0.15 [0.09,0.22]	0.20 [0.11,0.29]	0.04
	0.7 [0.63,0.76]	0.73 [0.69,0.78]	0.55

TABLE 3.9: Study 1: Coefficient estimates ( ) for each of the predictors in the best fitting model in Table 3, along with the 95% HDI. : Correlation (plus 95% HDI) with search length. R^2: Proportion of variance in search length that is predictable from the independent variable

Models	B_{m0}	B_{mf}
+ +	$1.9x10^{25}$	1
+	$1.6x10^{25}$	0.8
+	$3x10^{20}$	$1.5x10^{5}$
+	$1.02x10^{20}$	$5x10^{6}$
	0.64	$3.3x10^{26}$
	0.26	$1.3x10^{26}$

TABLE 3.10: Study 2: N=5: Ranking of the different models for the analysis of the search length using Bayes Factors. Selected model is in bold. The six models with the highest support are presented.

Parameters			R^2
intercept	2.97 2.94, 3.01		
	-0.08[-0.09,-0.07]	-0.87 [-0.92,-79]	0.76
	-0.03 [-0.06,-0.00]	0.16 [-0.07,0.4]	0.03
	1.36 [0.88,0.82]	0.17 [-0.06,0.4]	0.03

TABLE 3.11: Study 2: N=5: Mean of the posterior distribution of the covariates from the the model with the highest support. Each parameter and its correlation with search length is associated with the lower and upper limits of 95% highest density interval

Models	B_{m0}	B_{mf}
	$1.9x10^4$	1
+ ^{2}	$6.4x10^3$	0.32
+	$4.6x10^3$	0.23
+	$4.1x10^3$	0.21
+ ^{2} +	$2.3x103$	0.12
+ +	$1.6x10^3$	0.08

TABLE 3.12: Study 2: N=5: Ranking of the different models for the analysis of the performance using Bayes Factors. Selected model is in bold. The six models with the highest support are presented.

Parameters			R^2
intercept	27 26.7, 27.3		
	0.7[-0.42,0.98]	0.53[0.35,0.67]	0.28

TABLE 3.13: Study 2: N=5: Mean of the posterior distribution of the covariates from the the model with the highest support. Each parameter and its correlation with search length is associated with the lower and upper limits of 95% highest density interval

Models	B_{m0}	B_{mf}
+	$4.2x10^{28}$	1
+ +	$5.2x10^{27}$	0.13
	$1.4x10^{24}$	$3.3x10^{5}$
+	$1.4x10^{23}$	$3.3x10^{6}$
	1.46	$3.5x10^{29}$
	1.01	$2.4x10^{29}$

TABLE 3.14: Study 2: N=10: Ranking of the different models for the analysis of the search length using Bayes Factors. Selected model is in bold. Only the five models with the highest support are presented.

Parameters			R^2
intercept	4.7 4.6,4.81		
	-0.14[-0.16,-0.13]	-0.90 [-0.93,-0.85]	0.81
	3.06 [2.0,4.1]	0.22 [0.00,0.43]	0.05

TABLE 3.15: Study 2: N=10: Mean of the posterior distribution of the covariates from the the model with the highest support. Each parameter and its correlation with search length is associated with the lower and upper limits of 95% highest density interval

Models	B_{m0}	B_{mf}
+ ^{2} +	$4.0x10^{16}$	1
+ ^{2} + +	$5.4x10^{15}$	0.13
^{2} +	$1.1x10^{15}$	0.03
^{2} + +	$1.7x10^{14}$	0.004
+	$8.4x10^{12}$	$2.0x10\ ^{4}$
+ +	$1.2x10^{12}$	$2.6x10\ ^{5}$

TABLE 3.16: Study 2: N=10: Ranking of the different models for the analysis of the performance using Bayes Factors. Selected model is in bold. Only the five models with the highest support are presented.

Parameters			R^2
intercept	27.5 27.2,27.8		
	-0.55[-0.88,-0.24]	-0.30 [-0.50,-0.07]	0.09
^{2}	-0.54[-0.76,-0.32]	-0.67 [-0.79,-0.53]	0.45
	1.36 [0.99,1.7]	0.71 [0.58,0.81]	0.50

TABLE 3.17: Study 2: N=10: Mean of the posterior distribution of the covariates from the the model with the highest support. Each parameter and its correlation with search performance is associated with the lower and upper limits of 95% highest density interval

Models	B_{m0}	B_{mf}
+	$9.1x10^{18}$	1
+ +	$9.5x10^{17}$	0.10
	$6.1x10^{15}$	0.0
+	$6.0x10^{14}$	$6.5x10\ ^{5}$
	0.5	$5.11x10\ ^{20}$
	0.3	$3.2x10\ ^{20}$

TABLE 3.18: N=20: Ranking of the different models for the analysis of the search length using Bayes Factors. Selected model is in bold. Only the five models with the highest support are presented.

Parameters			R^2
intercept	7.4 7.11,7.68		
	-0.25[-0.28,-0.21]	-0.81 [-0.88,-72]	0.66
	8.9 [5.2,12.7]	0.06 [-0.16,0.29]	0.004

TABLE 3.19: N=20: Mean of the posterior distribution of the covariates from the the model with the highest support. Each parameter and its correlation with search length is associated with the lower and upper limits of 95% highest density interval

Models	B_{m0}	B_{mf}
+ **2** +	$1.4x10^{20}$	1
+ ^{2} + +	$3.1x10^{19}$	0.23
^{2} +	$3.3x10^{18}$	0.02
^{2} + +	$1.6x10^{18}$	0.01
+	$1.3x10^{11}$	$9.3x10^{10}$
+ +	$2.2^{1}0^{10}$	$1.6x10$ ^{10}

TABLE 3.20: N=20: Ranking of the different models for the analysis of the performance using Bayes Factors. Selected model is in bold. Only the five models with the highest support are presented.

Parameters			R^2
intercept	727.6 27.3,27.8		
	-0.44[-0.68,-0.19]	-0.07 [-0.3,-0.16]	0.005
^{2}	-0.63 [-0.78,-0.47]	-0.66 [-0.78,-0.51]	0.44
	1.6 [1.08,1.62]	0.70 [0.57,0.81]	0.49

TABLE 3.21: N=20: Mean of the posterior distribution of the covariates from the the model with the highest support. Each parameter and its correlation with search length is associated with the lower and upper limits of 95% highest density interval

SEQUENTIAL SEARCH TASK: HOW ITS COMPONENTS AFFECT DECISION STRATEGIES AND RISK PREFERENCES

Authors Christiane Baumann, René Schlegelmilch and Bettina von Helversen

Status working paper

ABSTRACT

Sequential decision making – a decision where available options are encountered successively - is a hallmark of our everyday life. When we search for a job or an apartment, we may decide to accept or reject it without knowing potential future options. Consequently, choices in such tasks correspond to risky decisions between a certain outcome and the risk of continuing search. However, risk preferences assessed in comparison to the standard optimal model do not correlate with indices elicited in single risky decision tasks. Moreover, humans make choices as if they are risk averse in the beginning but risk seeking at the end of search. This paper examines the influence of the characteristic components of sequential search tasks on people's search behavior which may give rise to the observed risk inconsistencies across the course of search. Results show that the sequential order and the unequal frequency in accept and reject decision, but not the presentation format of the

underlying distribution, leads to a systematic bias in people's choices. We conclude that the characteristic environment of a sequential search task takes precedence over stable individual risk preferences which is in part the reason for lack of convergence thereof between dynamic and single risk taking tasks.

4.1 INTRODUCTION

Whereas in classic economic decisions, individuals choose from a menu of well-defined options presented simultaneously, in many real-world decisions the options are encountered serially and cannot directly be compared to one another. For example, when we search for a job, a partner or the perfect day to open a good bottle of wine, we have to choose an option without knowing potential future options. Individuals, consumers and firms encounter many variations of these sort of sequential search problems and they have been well studied in behavioral psychology (Rapoport et al., 1966, 1970), marketing science (Moorthy et al., 1997; Zwick et al., 2003) and economic theory (Stigler, 1961). The difficulty in optimal stopping problems lies in the dynamic trade off between accepting a sub–optimal option too early or to reject the right option out of false hopes for a better one in the future. Existing research finds that people are very heterogeneous with respect to their behavior in optimal stopping situations (Guan et al., 2018; Guan et al., 2020; Lee et al., 2004), and it has been suggested that this heterogeneity in dynamic choice situations could be reflected in risk preference heterogeneity (Cox et al., 1989; Schotter et al., 1981). Consequently, many authors have proposed that the common finding of stopping too early in such tasks can be attributed to peoples general tendency of risk aversion (in the gain domain) (Schotter et al., 1981; Schunk et al., 2009; Sonnemans, 1998).

However, studies attempting to link individual risk preferences elicited in single gamble tasks with search behavior in sequential decision tasks have failed to show any relationship (Frey et al., 2017; Pedroni et al., 2017; Schunk, 2009; Schunk et al., 2009). It has been argued that the lack of convergence stems from the specific methods of the chosen elicitation methods which alter the way individuals evaluate options (Pedroni et al., 2017). Some go even further by assuming that the specific search environment overshadows any general preferences for exploitative behavior thus predictors for search behavior are exclusively restricted to the particular environment (Meyers et al., 2020). Indeed, it appears that human choice behavior on sequential tasks is confounded by complex interdependence between cognitive, motivational, and response processes, making it difficult to sort out and identify the specific processes responsible for the observed behavior (Brehmer, 1992; Busemeyer et al., 2002; Edwards, 1962; Gonzalez et al., 2017). However, while there is an enormous experimental literature on the foundations of decision behavior in static decision situations, the foundations of behavior in dynamic decision situations, despite being equally important, remain largely unexplored.

The standard optimal search model assumes expected value maximization (i.e. risk neutrality) such that an option is accepted if it is above a decision threshold that corresponds to the expected reward of the remaining options (see next section for a detailed description of the optimal solution). Accordingly, a risk averse subject would use lower than optimal decision thresholds and thus stop at options that are below the expected value of the remaining options. This would, in turn, lead to a shorter search length than prescribed from the optimal model. However, a recent study (Baumann et al., 2020) demonstrated that while participants tend to stop too early in the beginning of the sequence, they accept too little in the later phase of their search. In particular, participants revealed a risk seeking behavior on the second–to–last position, on which participants choose to either accept the safe option or draw one last time from the sampling distribution. Under the assumption of a symmetric sampling distribution, a rational and thus risk neutral strategy would lead to an acceptance rate of 50%, accepting options above the expected mean of the sampling distribution and rejecting options below. However, participants acceptance rates were significantly lower, accepting only 29% (95%-CI: [26%,32%]) of the encountered options. This result indicates that the particular parts of a sequential search task may affect peoples' choices in such a way that it appears as if people reverse their risk preference across search.

The goal of this paper is to examine the characteristic components of optimal stopping task that have an effect on choice behavior and thus encourage the emergence of risk inconsistencies. To do this, we confront subjects with different versions of an optimal stopping problem which are identical from an economic perspective, in the sense that the probability distributions over outcomes are the same. Specifically, our two studies examine the influence of (1) the presentation format of the underlying distribution, (2) the sequential order and (3) the unequal frequency in accept and reject decision affect peoples' decision strategy during search.

The structure of this paper is as follows: First, we explain a typical optimal stopping problem and its optimal solution; second, we describe the details of our experiments and discuss their results.

4.1.1 *Optimal Stopping task and its optimal solution*

Peoples' risk preference is commonly assessed relative to a model that assumes expected value maximization which embodies risk neutrality (Bernoulli, 2011). A risk neutral agent is indifferent towards the choice between a certain value and a risky option which holds an equal expected value. Consequently, a risk-avers person prefers to choose the safe option in this scenario, whereas a risk seeking person tends to choose the risky one. Accordingly, the normative model in optimal stopping tasks assumes expected value maximization implying that the decision maker chooses the sampled option if its value exceeds a threshold which corresponds to the expected value of continuing search. Optimal thresholds are thus derived based on the reward distribution of the available options. Its calculation is described based on the following optimal stopping example:

We consider a decision maker who encounters a sequence of options with rewards denoted by x_N, ,x_1 and she wants to find the maximum value in the sequence. If she accepts option i, then the sequence terminates and she receives x_i; otherwise, she continues to the next option. When the last option 1 is reached, it must be accepted. The optimal policy is to choose option i when it goes above a position-dependent threshold T_i. The calculation of the optimal thresholds applied to our scenario, where payoff is proportional to the chosen value and the goal is to find highest valued option, is described as follows (see also Gilbert et al., 1966):

Let's assume that a sequence of N options is drawn from a standard normal distribution with density

$$f\ x \qquad \frac{1}{\overline{\ 2}} e^{\ \frac{1}{2}\frac{t}{\ }^{2}} \tag{4.1}$$

and the goal is to find the highest reward in this sequence.

The *optimum strategy* for this task is:

If N 1, the decision maker is forced to accept the option, therefore the threshold of the last ticket is:

$$T_1 \quad 0 \tag{4.2}$$

Consequently, the expected reward (R_1) corresponds to the mean ( ) of the distribution.

If N 2, the decision maker decides to either keep the option on the first position x_2 or to continue to the last one x_1. If she decides to continue, the expected reward of x_1 is . Therefore she keeps x_2, if x_2 , rejects it if x_2 and is indifferent if x_2 . Therefore, the expected reward of x_1 (R_1) is also the threshold on the second last position:

$$T_2 \quad R_1 \tag{4.3}$$

Than for N 2, the expected reward (R_2) is calculated as follows:

$$R_2 \qquad \int_{R_1} f\ x\ \ x dx \quad R_1 \int^{R_1} f\ x\ dx \tag{4.4}$$

The remaining terms of the sequence can be computed in a recursive manner. For each N, the decision maker accepts the option if it is above the expected value of the remaining $N \quad 1$ options ($x \quad R_{n\ \ 1}$) and reject it otherwise ($x \quad R_{n\ \ 1}$). Therefore the threshold on the n-th position (T_n) is:

$$T_n \quad R_{n\ \ 1} \tag{4.5}$$

Accordingly the expected value of the reward (P_n) is :

$$R_n \qquad {}_{R_{n\ \ 1}} f\ \ x \quad xdx \quad R_{n\ \ 1} \qquad {}^{P_{n\ \ 1}} f\ \ x\ \ dx \tag{4.6}$$

Figure 4.1 shows the monotonically decreasing thresholds which correspond to the optimal solution in our studies with options sampled from 160, 20 . Additionally, the sampling distribution is displayed on each position indicating the range of options that are accepted on each position. It shows that the monotonically decreasing threshold involves an adaptation of the acceptance rate across position. Whereas only 12% of the encountered options are accepted on the very first position, the acceptance rate increases to 50% on the second-to-last position, indicated by a threshold that corresponds to the mean of the sampling distribution (here $T_2 \quad 160$).

In this sense, subjects using lower than optimal thresholds and thus accept more often are classified as risk averse. In contrast, risk seeking subjects set their threshold higher, thus becoming more selective.

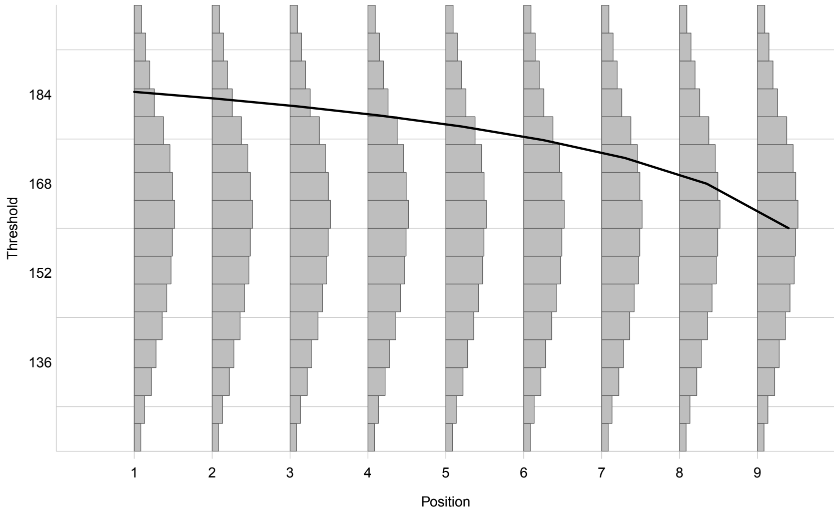

FIGURE 4.1: Optimal thresholds for the optimal stopping task with 10 options sampled from 160, 20 . Black line represents the decreasing optimal thresholds across position. Histogramms illustrate the sampling distribution of the safe option 160, 20

4.1.2 *Presentation of sampling distribution*

To decide when to stop searching depends largely on the sampling distribution of the alternatives. Therefore, distortions in the assessment of the distribution's estimates may lead to systematic biases in peoples' decision thresholds. In an optimal stopping task, subjects discover the sampling distribution by experiencing options from the respective distribution (here 160, 20). Several studies have shown that the way how we acquire information about distribution can lead to large behavioral differences (Barron et al., 2003; Erev, 2012; Hertwig et al., 2009; E. Weber et al., 2004). In particular, experimental investigation of decision-making in humans relies on two distinct types of paradigms, involving either experience- or description-based choices. In experience-based choices, outcome distributions are inferred by experiencing repeated draws from the distribution, corresponding to the way how distributions are presented in the optimal stopping task. In decisions by description, outcome distributions are explicitly described, by numerically presenting outcomes and probabilities. Barron and Erev (2003) demonstrated that the deviations from maximization that one observes in choices between lotteries depend critically on how the information was acquired (i.e, through a description or through experience). Furthermore, it was suggested that outcome and probability information translate into systematically different subjective representations in description- versus experience-based choice. Following this line of research, we assume that biases in optimal stopping behavior may emerge from the format presentation of the underlying sampling distribution. Therefore, the goal of the first study is to test whether there are behavioral differences in experienced based versus description based decisions in optimal stopping problems. To do that, we designed an experiment where participants perform two optimal stopping tasks which only vary in the presentation of the sampling distribution. In one task, the underlying sampling distribution is discovered by encountering alternatives. In the other task, outcome and probabilities are numerically presented thus providing the full information about the underlying sampling distribution.

We hypothesize that the format of the presentation, by experience or by description, may cause different biases in the estimates of the underlying distributions which would result in differences in search behavior. Furthermore, the indication of the full information of the sampling distribution throughout search should increase the transparency of the dynamic task structure and accentuate the independence between options. Therefore we expect that the full information of the underlying distribution leads to less biased decision thresholds, leading to a mitigation of the apparent inconsistency in the risk preferences.

4.2 STUDY 1: EXPERIENCE VERSUS DESCRIPTION

Participants performed two optimal stopping tasks which varied in the way how the underlying sampling distribution was presented. In the first task, participants encountered values sampled from the distribution and thus inferred its estimates based on experience (*Experience Task*, ET). In the second task, the sampling distribution was numerically displayed, thus participants had full information about the task structure (*Descriptive Task*, DT). The goal of this study was to understand the effect of decisions by experience or by description on participants search behavior in optimal stopping tasks.

4.2.1 *Methods*

4.2.1.1 *Participants*

We recruited 70 participants (26 females; age range: 20-70) on Amazon Mechanical Turk to participate in the experiment. Participants gave informed consent, and the study design and methods were approved by the ethics committee of the University of Zurich. Participants received a fixed payment of $3 and a performance dependant bonus ranging between $0 - $4.

4.2.1.2 *Procedure*

Figure 4.2 schematically depicts the procedure of the two types of tasks. In the ET, participants were instructed to sell their crashed car to a scrapyard for the highest price possible. In each trial, they could visit up to ten offers. If they declined the offer, they went on to the next one and it was not possible to go back in time. The trial ended if they accepted an offer, or if they arrived at the 10th offer which they were forced to accept. Price offers were sampled from a normal distribution with a mean of 160 and a standard deviation of 20. The bonus was performance contingent, by randomly choosing one of the 60 trials and dividing the chosen offer by 10 (for example an offer accepted at a price of 170$ would result in a bonus of 1.70$). In order to avoid learning effects during testing, participants performed a learning phase in which they were exposed to a total of 60 price offers from the same distribution (160, 20 , see Supporting Material Figure 4.8 for the details about the learning procedure).

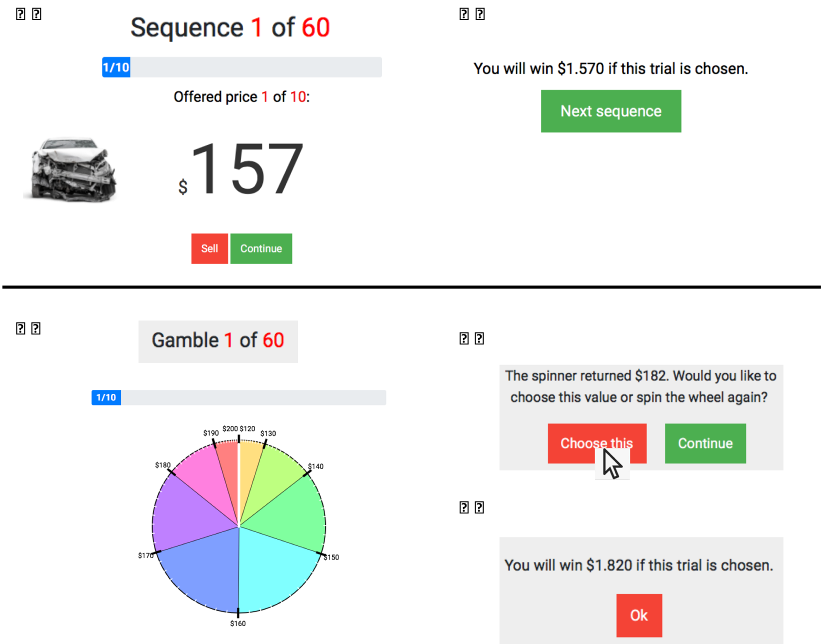

FIGURE 4.2: A1-A2: Screenshots of the experienced task (ET). The cover story is about finding the best offer for a broken car. In each trial, they could visit up to ten offers. B1-B3: Screenshots of the descriptive task (DT). People were instructed to win as much money as possible in a casino gambling task. In each trial, participants could spin a wheel of fortune up to 10 times and after each gamble they either accepted the outcome and proceeded to the next trial.

The DT imitated a casino gambling scenario where people tried to win as much money as possible (see Fig. 4.2 B1-B3). In each trial, participants could spin a wheel of fortune up to 10 times and after each gamble they either accepted the outcome and proceeded to the next trial, or rejected it and spun again. The wheel of fortune was not changing throughout the experiment and represented a normal distribution with a mean of 160 and a standard deviation of 20, equivalent to the sampling distribution in the ET. Participants were always aware of the total number of spins in each trial and of the actual position in the sequence (see Fig. 4.2 B1). It was not possible to go back to an earlier outcome of a gamble after it was initially declined. If they spun the last gamble (10*th*) they ware forced to choose the outcome. The bonus corresponded to the outcome of one randomly chosen trial divided by 100, following the same procedure as in the ET.

The two tasks were randomly presented and consisted of 60 trials. To ensure that participants encountered the exact same values in both tasks, we generated 60 sequences prior to the first task (60 x 10 values 160, 20). The order of the sequences in both tasks was randomized and each participant encountered a newly generated sample.

4.2.2 *Results*

4.2.2.1 *Search length and performance*

First we analyzed if the presentation manipulation influenced behavioral measurements such as search length and performance (see violin plots, Fig. 4.3 A and B). Participants average search length (measured as the number of 'reject' decisions in a sequence plus 1) was 4.7 in the ET (SD = 1.03) and 4.1 (SD = 0.86) in the DT, with 78% of the participants searching more in the ET than in the DT. The average accepted price was 180.3 (SD=2.0) in the ET and 180.1 (SD=2.7) in the DT. The optimal model without any error would search on average 5.1 and choose options with an average reward value of 182.5 suggesting that participants did not search in accordance with the noiseless optimal model in our experiment in both tasks. A Bayesian t test (prior scale = medium, Morey et al., 2018) provided evidence that participants search less in the DT compared to the ET (M_{Diff}^{SL} 0.61, HDI_{95} 0.33, 0.88, BF_{10} 300). The average accepted prices in both conditions differed only slightly, which is supported by a Bayesing t test showing moderate evidence for no difference in the mean accepted price between the ET and the DT (M_{Diff}^{SL} 0.17, HDI_{95} 0.25, 0.58, BF_{10} 0.2).

We find stable individual differences in search length and the average accepted price between the ET and DT, indicated by a correlation coefficient of $r = 0.7$ [0.53,0.82] between search length and r =0.74 [0.59, 0.84] between the accepted option value (BF_{10} 300 for both tests, analysis conducted using (prior scale = medium, Morey et al., 2018)).

To investigate whether the probability to accept differed between the ET and the DT across the sequence, we used a logistic mixed model (for an introduction, see (Singmann et al., 2019); see also (Bolker et al., 2009)) with the choice response as dependent variable (accept (1) vs reject (0)) and Condition (ET vs. DT), Position (a centred continuous variable denoting the 9 positions) and the interaction of Condition Position as fixed effects, and Participants as random intercepts as well as random slopes for Condition Position. The model revealed a significant effect of Condition, ^{2} 5 102.04, p .0001, whereas participants accept overall more in the DT than in the ET. The estimated marginal means (EMM) for the DT exhibited a higher acceptance probability: EMM_{ET} = 21%, 95%-CI [20%, 23%], EMM_{DT} = 25%, [23%, 26%]. Furthermore, the analysis of choice probabilities shows a significant increase across position (^{2} 5 914.83, p .0001) in both conditions and a significant effect of the interaction of Condition Position (^{2} 5 23.95, p .0001). As shown in Figure 4.3 C, participants in the DT accepted more often in the beginning of the sequence compared to the ET, but adapt their rate of acceptance less strongly across the sequence.

To get a better understanding of the divergence in acceptance probabilities between the two tasks, we calculated participants' acceptance probability dependent on the accepted values. To do so, we split the accepted values into quantile ranges, whereas Qi is defined as the range of values from the $0.ith$ to the $0.i$ 0.1 th quantile of the sampling distribution. Figure 4.3 D shows $Q6$ - $Q10$ (out of a total of ten quantile ranges). Thus it shows that differences in acceptance rates are mainly driven by values ranging in the upper 20% of the sampling distribution (Q8-Q10). In particular, acceptance rates for these values are higher in the DT than in the ET.

4.2.2.2 *Does the presentation format affect risk preference inconsistencies?*

We next address the question about the impact of the descriptive format on the decision strategy leading to reversals in risk preferences. We expected that the full information of the distribution leads to a more independent update of the decision thresholds across search, which in turn would result in an attenuation of the observed asymmetries in risky behavior. Participants' decision thresholds were calculated in both conditions (ET and DT) by fitting a logistic curve (i.e., generalised linear model; GLM) for each combination of participant and position

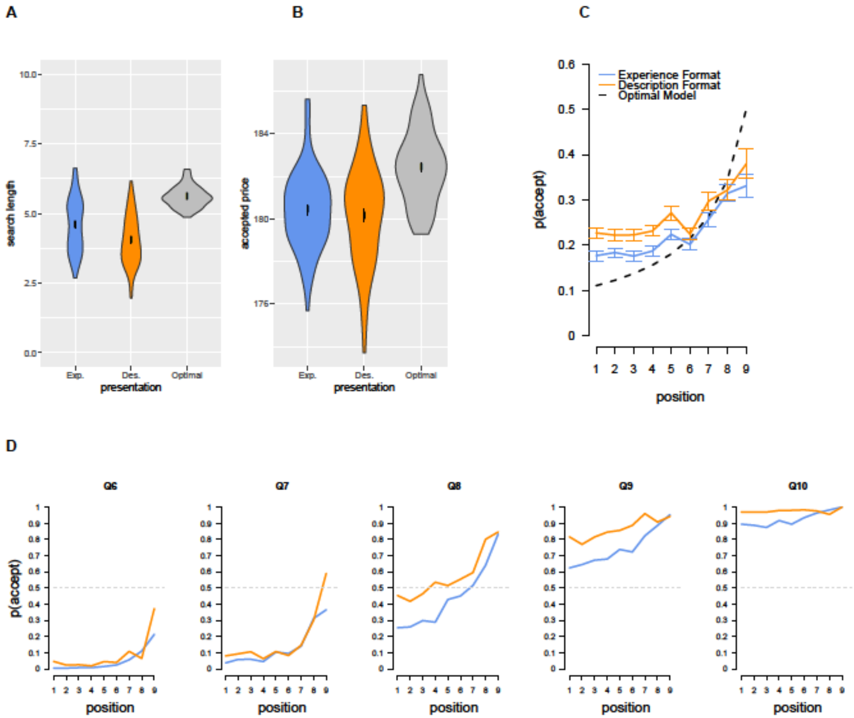

FIGURE 4.3: A: Violin plots of participants' search length in both presentation conditions blue: ET, orange: DT, grey: optimal. B: Violin plots of participants' average accepted price, blue: ET, orange: DT, grey: optimal search length if optimal policy is applied); C: probability to accept on each position, blue: sequential offer task, orange: sequential gambling task, black: optimal. bars: standard errors of the mean, D: Probability to accept an offers/certain outcomes divided into value ranges. Q6: Values ranging between the 5th and 6th quantile, Q7: Values ranging between the 6th and 7th quantile, etc.

(i.e., the total of number of GLMs was number of participants times 9). Each GLM was fitted to participants' choices (accept: 1 or reject: 0) as dependent variable with the independent variable being the certain outcome (in the DT condition) or the offer value (in the ET condition). From each GLM we calculated participants' individual decision thresholds (i.e., the indifference point between accept and reject) as minus intercept divided by the slope. The individual decision threshold defines the point on the continuous variable (i.e., value of the certain option/offer) at which the logistic function predicts an acceptance rate of 50%. Figure 4.4 displays participants average decision thresholds on each position (blue: ET, orange: DT, black: optimal thresholds), together with the individual decision thresholds. Indeed, 82% of the participants in the ET and 95% of the participants in the DT have a lower than optimal threshold in the beginning of their search, thus showing risk averse preferences. On the second-to-last position, 96% of the participants are above the optimal threshold in the ET and 98% in the DT, revealing risk seeking preferences. Consequently, the observed inconsistency of risk preferences in optimal stopping tasks remains in the DT, despite the full information about the probabilities and outcomes on each time step. This finding is in contrast to our hypothesis which stated that the descriptive format is more informative and thus leads to a more consistent adaptation of the decision thresholds compared to the optimal, i.e., risk neutral model.

Figure 4.4 B indicates participants average reject rates on each position in the ET, the DT and when following the optimal rule. Reject decisions are predominant in optimal stopping tasks in both task conditions, with an overall participants' reject rate of 76% in the ET and 72% in the DT.

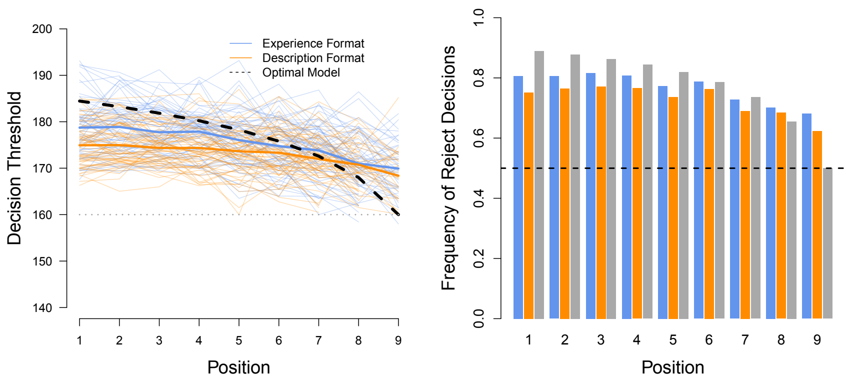

FIGURE 4.4: A: Individual (transparent lines) and average decision thresholds (bold lines) across positions (blue: ET, orange: DT). Optimal threshold: black dotted line. Decision thresholds represent the fixed value at which participants are indifferent to accept it or further search/gamble. Light gray dotted line represents the mean of the distribution. B: Frequency of reject decisions on each position. ET: blue bars, DT: orange bars, optimal: grey bars. Black dashed line indicates a rejecting rate of 50%.

4.2.3 *Discussion*

In an optimal stopping task, subjects discover the sampling distribution by experiencing options from the respective distribution. Thus we speculated that the presentation of the underlying sampling distribution, where samples are experienced, may bias choice behavior leading to apparent inconsistent risk behavior. Therefore, we contrasted the conventional optimal stopping task, which involves inferring the distribution by encountering values, with a gamble that provided the a description of outcomes and probabilities. We speculated that the descriptive format of the underlying distribution would lead to a difference in search behavior. In particular, providing the full information should enhance the transparency of the task, resulting in an attenuation of risk inconsistencies across search.

Our results indicate that the presentation format leads to significant differences on participants' choice behavior, whereas acceptance probabilities are higher when the distribution is described. This finding agrees with the hypothesis that the way how estimates are inferred, by experience or description, may lead to biases in their search behavior. From a risk preference perspective, people thus are more risk seeking in the experienced format throughout their search. This result is in line with the observation that participants become more risk seeking in the gain domain when decisions are based on experience (Ludvig et al., 2018). It has been argued that in decisions based on experience, extreme values (big wins or big losses) carry proportionally more weight, leading to higher risk seeking for gains and risk aversion for losses. Some evidence for this potential extreme-outcome rule comes from studies with non-human animals, which can only rely on experience for learning about outcomes. Many of these studies have also reported risk seeking for gains (Hayden et al., 2008; Heilbronner et al., 2013; Kacelnik et al., 1996; McCoy et al., 2005, e.g.), and some evidence suggests that this risk seeking for gains may be driven by extreme outcomes in a context. The contact with extreme outcomes in the experienced task might thus have elicited relatively more salient and weighted responses, driving the subsequent choices in their search for the optimal offer.

However, despite the difference in participants acceptance rates, the format of the presentation has a minor effect peoples' strategy of updating their decision thresholds. In particular, we observe the same pattern in both conditions compared to the optimal model: Whereas decision thresholds are too low in the beginning of search, they are too high at the final positions. Therefore, the asymmetry in risk preferences remain in both condition. This finding is surprising, since we had assumed that participants exposed to the full information will be less sensitive to the sequential character of the task and thus update their decision thresholds more independently. Especially on the second–to–last position, which corresponds to a single choice between a fixed outcome or spinning the wheel one more time, we would have expected that participants approach their acceptance rate more closely to 50%. On this position, the certain outcomes spread evenly around the mean, with 50% above and 50% below the mean.

A possible explanation for the similar biases in choice behavior in the descriptive format could be that feedback was provided after each choice. It has been suggested that in this description and experience case, information from description may be combined with information from experience (Jessup et al., 2008). Lejarraga (2011) argued that information from description is neglected in the presence of feedback. In that sense, the feedback provided in the descriptive format may override any benefit given by the full information, thus both presentation formats lead to same biases in participants choices.

We thus conclude that the presentation of the underlying sampling distribution has a minor impact on the decision strategy which causes the apparent inconsistencies in risk preferences across search. We rather speculate that the dynamic character of the task may trigger other cognitive mechanisms that override stable risk preferences. Indeed, research exploring dynamic effects has found that individuals become more risk-seeking following losses and argue that due to loss aversion, the individual becomes more willing to accept a lottery if it offers a possibility of erasing the previous negative outcome (Rabin et al., 2009; Read et al., 1999). In that sense, during the course of search, participants may substitute away from normative decision thresholds (thus expected values) towards outcomes that have a lower likelihood of appearing, in an effort to obtain the option they have hoped. We will address this issue in the second study.

A second explanation for increased risk seeking behavior at the end of the sequence is the overall predominance of reject decisions in optimal stopping tasks (see Figure 4.4 B). Participants arriving on the last positions may just tend to reject the certain outcome despite its superiority compared to the risky option. Erev and Haruvy (2016) review studies in which participants repeatedly choose between a risky prospect and a safe option, and receive immediate feedback (e.g., Erev, 2012). They (Erev et al., 2016) conclude that there exists a strong tendency to simply repeat the most recent decision, which is even stronger than the tendency to react optimally to the most recent outcome. In that sense, the inequality in accept and reject choices inherent to an optimal stopping problem may lead to a bias towards reject decisions on the later positions, which we will examine in the next study.

Sequential choices

An important characteristics of optimal stopping task is its dynamic aspect during search: It involves a sequence of choices where decisions are sequentially linked so that rejecting an option at a specific time directly influences future choices. Numerous studies have shown that decision with risky or uncertain outcomes are affected by the outcomes of previous decisions (Ofek et al., 2007; Thaler et al., 1990; M. Weber et al., 2005). Kahneman and Tversky (1979, p.286) also recognize that "there are (..) situations in which gains and losses are coded relative to an expectation or aspiration level that differs from the status quo". In these situations,"the outcomes of an act affect the balance in an account that was previously set up by a related action" (Tversky et al., 1981, p. 457). For example,"a person who has not made peace with his losses is likely to accept gambles that would be unacceptable to him otherwis" (Kahneman et al., 1979, p.287). In an earlier study, McGlothin (1956) has shown that a person betting on a long-shot at the end of the racing day would be fully aware of the "riskiness" of the horse but seek it out to recover his or her losses. Similarly, research exploring dynamic effects has found that individuals become more risk-seeking following losses and argue that due to loss aversion, individuals becomes more willing to accept a lottery if it offers a possibility of erasing the previous negative outcome (Rabin et al., 2009; Read et al., 1999). Rabin et al. (2009) conclude that losses are integrated and evaluated jointly with prospects in the same brackets. Imaz (2016) has shown that the increase in risk–taking following a paper loss is a product of dynamic inconsistency in preferences – individuals deviate from their planned risk-taking strategies to take on more risk after a paper loss, and that realization of a loss mitigates these deviations.

Following these lines of research, we suggest that participants' risk seeking behavior at the later stages in search might be caused by a feeling of hope that the anticipated option must appear in an additional search. As a consequence, participants update their decision thresholds depending on earlier outcomes and not – as prescribed by the optimal policy – exclusively on the remaining future options. In the second study, we aim to understand if inconsistencies in risk preferences might emerge from the dynamic structure in optimal stopping tasks. To do this, we contrasted search behavior in an optimal stopping task with choices in single gambles with identical statistical properties. We hypothesise that removing the sequential character between choices leads to independent updating of decision thresholds between gambles. This, in turn, will mitigate the inconsistency in risk preferences found in optimal stopping tasks

Imbalance of accept and reject decision

A further important characteristic in optimal stopping tasks involves that only one option within the sequence can be accepted, which implies that all the options before have to be denied. Consequently, reject decisions are overrepresented throughout search (see also Fig. 4.4 for reject rate in Study 1). Whereas choices on early positions (position 1 - 6) should be rejected, according to the optimal, in up to 85%, the reject rate should be rigorously adapted on later positions, reaching an reject level of 50% at the second to last position.

However, people insist in rejecting sure outcomes on later positions, as if they become more risk seeking. We speculate that the apparent increase in risk seeking behaviour emerges from a bias to repeat prior choices. Studies have shown that people are prone to replicate their choices, no matter whether these choices led to success or failure (Charness et al., 2005; Erev et al., 2016). Consequently, removing the imbalance of accept and reject decisions should attenuate the tendency to repeat the reject decision on later positions, and thus reduce the increase in risk seeking behavior. To test our assumption, we added a third task which removed the inequality between accept and reject decisions.

4.3 STUDY 2: SEQUENTIAL PRESENTATION AND IMBALANCE IN CHOICES

In the second study, we examine the differential effect on behavior between optimal stopping choices and single gambles that are identical in outcome and probabilities. Furthermore, we test if the imbalance of accept and reject decisions, which is an inherent characteristic of optimal stopping tasks, has an additional influence on participants choices.

4.3.1 *Methods*

4.3.1.1 *Participants*

We recruited 70 participants (26 females; age range: 20-70) on Amazon Mechanical Turk to participate in the experiment. Participants gave informed consent, and the study design and methods were approved by the ethics committee of the University of Zurich. Participants received a fixed payment of $3 and a performance dependant bonus ranging between $0 - $4.

4.3.1.2 *Procedure*

All participants engaged in three tasks. One of these tasks was identical to the car selling task in Study 1. Participants had to find the best offer within a sequence of 10 sequentially presented offers in a total of 40 trials (*Sequential Offer Task*, SOT). Values were sampled from 160, 20 . Prior to the task, participants encountered a minimum of 60 values drawn from the sampling distribution (see

Supporting Material Fig. 4.8 for a description of the learning phase). In the second task, participants encountered 160 single gamble trials, in which they either could choose a certain outcome or to spin a wheel of fortune which displayed the outcomes and probabilities (*Single Gamble Task*, SGT, see Fig. 4.5). Importantly, choices were matched corresponding to the choices in the SOT on the $3rd$, $5th$, $7th$ and $9th$ position, by adjusting the expected value of the gamble to the expected value of search in the SOT. Outcomes and probabilities were adjusted accordingly, by determining the probability for specific value ranges. For example, if the participant encounters a safe value of 155 on position 3 in the SOT, he would encounter the same choice in the SGT in which he decides between accepting 155 or spin the wheel which corresponds to the outcomes and probabilities when arriving at the $3rd$ position (Fig. 4.5 A). This way, we could ensure that the gambles in the SGT have the identical statistical properties as the choices in the SOT (see Supporting Material Text 1 for a detailed description on the calculations of outcomes and probabilities for each wheel). Fig. 4.5 A - D shows the four wheels which display outcomes and probabilities according to the four positions in the SOT. Accordingly, the 160 trials consisted of 40 (trials) 4 (positions) (40 sequences from the SOT x 4 positions) which were randomly presented. We note that in the SOT, participants might have stopped earlier and thus have not encountered all the offers up to position 9, nevertheless, in the SGT, participants encountered all possible 160 gambles (4 wheels x 40 sequences = 160 trials). Consequently, we generated the 40 sequences before the first task started and used the exact same values in both tasks. The order of the 160 gambles was randomized and we introduced two trial runs prior to the real task.

In the third task was designed in the same fashion as the Holt and Laury gambles (Holt et al., 2002), where people decided between a certain outcome and spinning the wheel. The four wheels corresponded to the ones used in the SGT (see Fig. 4.5 A - D), however this time, the certain outcome ranged closely around the mean (mean +/- 5,10,15,20) with a total of 8 values for each wheel (we call it *short Single Gamble Task*, sSGT). This task always started with the gamble showing the wheel representing the 3rd position (see Fig. 4.5 A) and it changed after 8 trials to a gamble showing a wheel that represents the next position (here: position 5). Fixed outcomes were monotonically increasing, starting from the lowest value. Participants were instructed that the fixed outcome increased in each sequence of 8 trials and that the wheel changed after each 8 trials. There was a total of 32 trials.

In order to avoid that the two gambling tasks appeared by immediate succession, we arranged the order in the following way: The SOT appeared always on the second place, whereas the SGT and the sSGT were randomly assigned to the first or third place.

4.3.2 *Results: Sequential Presentation*

4.3.2.1 *Agreement with Study 1*

We first test if the behavioral measures in the SOT coincide with the identical sequential offer task in Study 1. Participants search length is on average 4.8 (SD = 0.9) and the average price is 180.5 (SD = 2.5), thus behavioral measures coincide with the measures obtained in Study 1 (ET). A Bayesian *t* test (R BayesFactor::ttestBF package, prior scale = medium; Morey et al., 2018) indicates no evidence for a

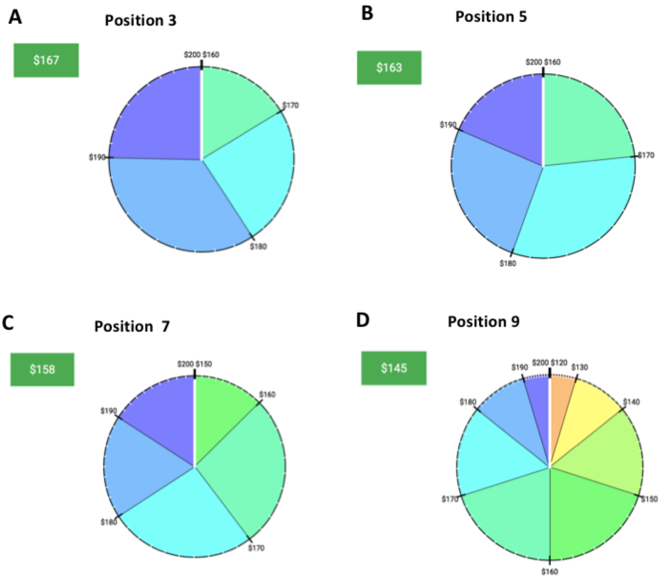

FIGURE 4.5: Study 2: Four wheels used in SGT and sSGT: Each wheel corresponded to the outcome distributions of the remaining options when choosing on the 3*rd*, 5*th*, 7*th* or 9*th* position in the SOT. Participants could either choose the certain outcome or spin the wheel with the hope for a greater reward.

difference in search length nor performance between the ET from Study 1 and the SOT in Study 2 (M^{SL}_{Diff} 0.3, HDI_{95} 0.62, 0.03 , BF_{10} 0.8, M^{Perf}_{Diff} 0.23, HDI_{95} 0.98, 0.48 , BF_{10} 0.22).

4.3.2.2 *SOT versus SGT*

Our first analysis examines if choice behavior differs between the sequential offer task and independent single gamble choices with identical statistical properties. Therefore, we compared participants' probability to accept on the $3rd$, $5th$, $7th$ and $9th$ position between the SOT and the SGT (Figure 4.6 A). For this analysis, we selected only trials in the SGT which were actually encountered in the SOT. For example, if a participant stopped at the $6th$ position in a sequence in the SOT, he encountered the $3rd$ and $5th$ position, thus only the trials that corresponded to these two choices were considered in the SGT. Visual inspection reveals that participants differ in their choices on the corresponding positions between the SOT and the SGT. Whereas in the SOT, participants increase their accept rates across position in a linear manner, in the SGT they seem to be more adapted to the optimal model. In order to test if choices differ between the two conditions, we used a logistic mixed model with the response variable (accept safe outcome (0) vs. reject/choose the gamble (1)) as dependent variable and Condition (SOT vs. SGT) and Position (as factor) and Condition Position as fixed effects (Participants as random intercepts as well as random slopes for Condition and Position). We used Type 3 likelihood-ratio tests (LRTs; in R afex::mixed; Singmann et al., 2019) which provide p values for nonzero differences inexplained variance between the full model (i.e., with all possible effects) and the restricted models (i.e., without the tested fixed effects). We find a significant main effect of Condition (^{2} 21 5.30, p 0.021) such that the estimated marginal means (EMM) for the SOT and the SGT exhibited an overall higher acceptance rate in the SOT: EMM_{SOT} = 22.3%, 95%-CI [20%, 24%], EMM_{DT} = 19.6%, [17%, 21%]. Furthermore, both Position (^{2} 21   78.24, p .0001), and the interaction Condition   Position (^{2} 5   10.97, p .01) reveal a significant effect. As shown in Figure 4.6 A, while in both conditions accept probabilities are increasing, participants accept less on early positions in the SGT but adjust their acceptance rate more strongly in gambles representing the 9*th* position. This result indicates that despite identical statistical properties, choice behavior differs between independent gambles and sequentially presented choices.

We further wanted to understand if participants' accept probabilities between the two tasks are consistent between the SOT and the SGT on the corresponding positions. A correlation analysis revealed moderate evidence for a correlations on the 3*rd* position of   = 0.27 [0.04,0.47] (BF=4.5) and strong evidence for correlations on the 5*th* (= 0.38 [0.17,0.57], BF=96), the 7*th* (= 0.39 [0.17,0.57], BF=93) and the 9*th* position (= 0.54 [0.36,0.69], BF>300, analysis conducted using (Morey et al., 2018, prior scale = medium).

Participant's risk preference has been shown to shift from risk averse to risk seeking during search in an optimal stopping task. Whereas decision thresholds are lower than in the normative model in the beginning of search, they are higher at the later stages of search. In order to examine if independent gambles with identical statistical properties lead to the same inconsistency in risky preferences, we determined the decision thresholds in both tasks by fitting a logistic function (i.e., GLM) to each participant's choice data (accept (1) vs. reject/gamble (0)) as a function of the fixed outcome. Figure 4.6 B (blue: SOT, orange: SGT) displays the decision threshold in both the SOT (blue line) and the SGT (orange line) and

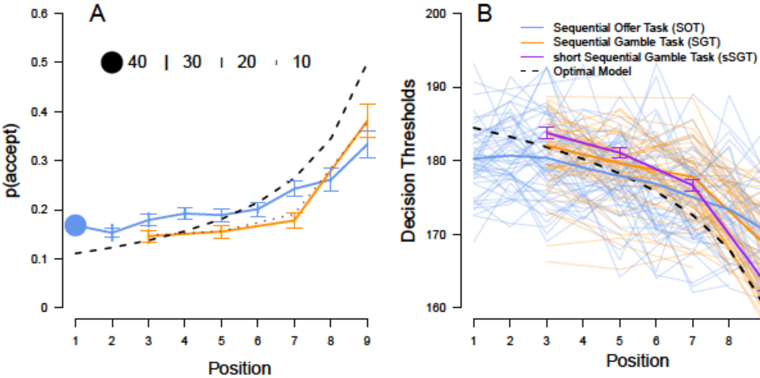

FIGURE 4.6: A: Probability to accept on each position, blue solid line: SOT, orange solid line: SGT (only choices included that correspond to choices in the SOT), dark orange dotted line: SGT (all choices), black dashed line: optimal model. The size of the dots represent the amount of data points per participant (see legend), error bars represent standard errors. B: Individual (transparent lines) and average decision thresholds (bold lines) across positions (blue: SOT, orange: SGT, purple: sSGT). Optimal threshold: black dotted line. Decision thresholds represent the fixed value at which participants are indifferent to accept it or further search/gamble.

indicates that that participants update their decision thresholds in closer agreement with the optimal model. Moreover, the reversal of risk preferences is reduced, thus suggesting that the sequential order of choices leads to decision strategies that override stable individual risk preferences.

However, participants in the SGT still exhibit an increase in risk seeking behavior in gambles representing the later positions in the SOT. The average thresholds on the $9th$ position are lower in the SGT than in the SOT (see Figure 4.6 B), however the comparison of the decision thresholds on the $9th$ position between the SOT and the SGT do not seem to differ (Bayesian t test, M_{Diff} 0.9, HDI_{95} 2.2, 0.58 , BF_{10} 0.2). However, both lie significantly above the optimal threshold (BF>300) which reflects peoples' apparent risk seeking preference. Consequently, despite the closer adaption of thresholds to optimality in single gamble tasks compared to sequential choices, the increased risk seeking behavior in gambles representing the later positions remains.

One possibility for this finding is the selective sample of choices due to the nature of optimal stopping tasks. It might be possible that risk averse participants accept earlier and therefore choices from the more risk seeking participants dominate on the final positions. In the above analysis, we only considered choices in the the SGT which were encountered in the SOT. In order to test this assumption, we compared the accept probabilities of the selective sample from the SGT, which corresponds to the choices encountered in the SOT, with the complete set of choices in the SGT, that is 40 choices on position 3, 5, 7 and 9, in Figure 4.6 A (orange line: selective data set, dark orange dotted line: full data set). Even a quick glance shows that the accept probabilities are essentially on top of each other, providing evidence that the selective sample in the optimal stopping task represents unbiased choice behavior on each position. A logistic mixed model with response (accept (0), gamble (1)) as dependent variable and Condition (all choices in the SGT versus only choices which were encountered in the SOT) and Position and the interaction of Condition Position as fixed effect and subjects as random intercepts and slopes supported the finding, showing no significant effect of Condition, ^{2} 21 0.1, p 0.9 nor its interaction with Position ^{2} 21 0.28, p 0.9.

4.3.3 *Results: Imbalance in Choices*

The inequality of reject and accept decisions inherent to optimal stopping tasks remained in the SGT, thus reject choices were overrepresented, most pronounced in gambles mirroring early positions in the SOT. We expect that the seeming increases of risk seeking behavior on later positions is partially caused by the tendency to repeat prior choices, which are predominately reject decisions. Therefore, we introduced a third task in which the unequal frequency of choices was removed. In the sSGT, the gambles corresponded to the four wheels of fortune presented in the SGT (Figure 4.5 A-D). Within each of the four gambles the certain option was initially lower than the expected value of the gamble (expected value - 15) and increased in steps of 5 (until expected value + 15), making the safe option more and more attractive. We assessed a person's decision threshold for each wheel by determining the point at which he switches from rejecting the safe option to accepting it. For instance, rejecting a safe option of 155 in favour of spinning the wheel, but accepting the following value of 160 would lead to a decision threshold of 157.7 ($option_{rejected}$ $option_{accepted}/2$). Fig. 4.6 B (purple lines) shows participants' average decision thresholds across the four gambles. It indicates that in contrast to the SGT, participants update their decision thresholds in close agreement with the optimal model. However, risk preferences are in closer agreement with the SGT, revealing an overall risk seeking behavior.

In order to measure stability in decision thresholds across the two tasks, we calculated correlations of decision thresholds between the SGT and the sSGT on each of position. We find very strong evidence for a correlation of $r = 0.47$ [0.24,0.66] (BF>100) and $r = 0.41$ [0.18,0.62] (BF=55) on gambles representing position 3 and 5 respectively, but no evidence for gambles representing position 7 and 9 (BF=0.33 and 0.34, respectively). This result suggests that the imbalance of reject and accept choices inherent in the SGT affects mainly choices on later stages of search.

In order to assess an absolute measure of participants' risk preference in the SGT and the sSGT, we subtracted the expected value of the corresponding gamble from participants decision thresholds, with positive values indicating risk seeking and negative values risk averse participants (see Figure 4.7). A separate test for both the SGT and the sSGT with absolute risk preference as dependent variable and position as independent variable (R BayesFactor::lmBF package Morey et al., 2018) provides strong evidence for an effect of position on risk preferences in the SGT (BF>300) but moderate evidence for no effect in the sSGT (BF=0.23). This result provides evidence that removing the imbalance of accept and reject choices attenuates the apparent risk seeking behavior on later positions, resulting in stable risk preferences across positions.

Furthermore, we were interested if risk preferences measured in the sSGT predict search behavior in the optimal stopping task. We used a mixed model with response (accept: 1, reject: 0) as the dependent variable and the average indifference point measured in the sSGT as fixed effect and Participants as random intercepts as well as random slopes for the indifference point. We find a significant effect of the indifference point on search behavior (^{2} 5 3.9, p .02), such that higher risk aversion measured in the gambling task leads to higher acceptance rates (and thus earlier stopping) in the SOT.

4.3.3.1 *Risk preferences in minimal context*

In contrast to the well documented phenomena that people tend to be risk averse in the gain frame (Kahneman et al., 2013a, 2013b; Tversky et al., 1981), participants performing the sSGT and the SGT revealed a general risk seeking preference relative to the optimal model across all gambles. These findings suggests that our particular task in which people spin a wheel of fortune, evokes a relatively higher risk seeking behavior, whereas safe options that exceeded the expected reward of the gamble were rejected. In order to confirm that participants are generally risk seeking in our tasks, we added a short third study including 50 participants on Amazon Mechanical Turk (20 females; age range: 23-69) in which we removed any additional context. The task was constructed so that the participant could either choose a certain outcome or spin a wheel of fortune. The wheel corresponded solely to the 9*th* position in the sSGT (see Figure 4.5 D), which represents the outcomes and probabilities of a normal distribution with mean of 160 and standard deviation of 20. A total of 20 safe outcomes were sampled from the same distribution, by making sure that 10 values were above and 10 below the expected value of the gamble.

We calculated an absolute measure of participants' risk preferences by fitting a logistic curve (i.e., GLM) to participants' choices (accept:1 or reject:0), with the difference between the safe outcome and the expected value of the gamble as the

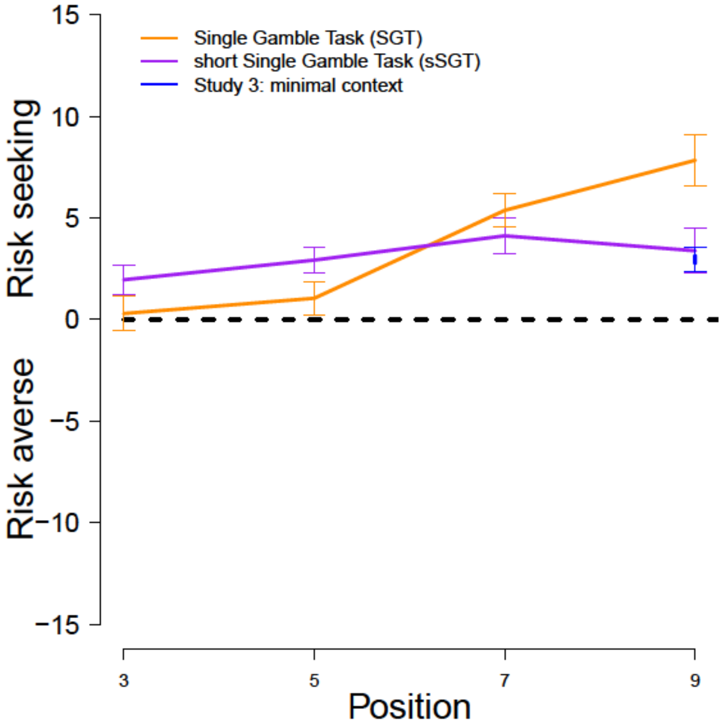

FIGURE 4.7: Average estimated indifference points (SGT: orange, sSGT: purple, Risk preference in minimal context: blue). sSGT on each of the measured position (3, 5, 7 and 9). Values above the black line indicate risk seeking behavior, thus preferring to gamble when safe outcomes have higher value than the expected value of the gamble. bars: standard errors of the mean.

dependent variable. Participants' average indifference point is shown in Figure 4.7 (in blue) on the 9*th* position. The result confirms the the basic trend of risk seeking preference gambling tasks. The comparison of the indifference points between the two tasks (new study and sSGT, on 9*th* position) provides strong evidence that risk preferences do not differ between these two data sets (Bayesian t test, BF=0.16). This result rules out that risk preferences in the sSGT are affected by further contextual effects and confirms a general risk seeking tendency in our wheel of fortune task.

4.3.4 *Discussion*

The first purpose of this study was to investigate the effect of two important characteristics inherent in optimal stopping problems on peoples' decision strategies and thus on the apparent emergence of risky inconsistent preferences. First, we investigated if the sequential character leads to a bias in participants decision thresholds, by contrasting choice behavior in an optimal stopping search and in single gambles which were identical in outcomes and probabilities. We find that decision behavior differs significantly between the conditions, whereas single gambles lead to an attenuation of the risk preference inconsistency found in optimal stopping problems. However, we still observe a seemingly increase of risk seeking behavior in gambles representing the later position in the optimal stopping task.

The second purpose of this study was to examine the effect of unequal proportions of reject and accept decisions on human's decision thresholds. We speculated that the tendency to repeat predominant reject decisions could lead to choices which look like as if people become more risk seeking at the later stage. Our results show that choices differ significantly in gambles with an equal proportion of accept and reject choices. In contrast to the increase in risk seeking behavior found in imbalanced choice situation, risk preferences are stable once this imbalance is removed. Furthermore, we found a general tendency of risk seeking behavior in the wheel of fortune task, which was confirmed in an additional study measuring risk preferences in a minimal context.

4.3.4.1 *Sequential vs single gamble decisions*

We have shown that choices in a sequential context deviate significantly from independent single gambles despite identical statistical properties. The reversal in risk preferences is attenuated in single gambles which represent the aggregated outcome distributions of the corresponding choices in the optimal stopping task. We speculated that the sequential character of the task may let participants think that if they only search enough, they will find a good option. The adherence to an aspiration level set prior to search leads to most pronounced deviations from optimality on the last positions, since these stages of search require to update the decision thresholds most strongly. Our finding is in line with a well documented phenomenon in repeated risky decisions, whereas Kahneman and Tversky (1979) recognized that "person who has not made peace with his losses is likely to accept gambles that would be unacceptable to him otherwise ".

4.3.4.2 *Effect of unequal frequency of accept and reject decisions*

Our results have shown that the overrepresentation of reject decision in optimal stopping tasks leads to an adherence to the same choice on later stages of the search.

Several studies have shown that decision inertia plays a role in human decision making under risk and that humans have the tendency to repeat previous choices in decision making with monetary feedback (Alós-Ferrer et al., 2016; Erev, 2012).

The imbalance of reject decisions implies that the range of values considered as losses is much wider than the range of values in the gain domain (see 4.1 on each expect on the 9*th* position. Several studies have shown that the set of available options in decisions under risk can affect the selection of the certainty equivalent, that is in our case, the decision threshold (Birnbaum, 1992; Stewart et al., 2003). Results of Walasek and Stewart (2015) have shown that a difference in the distributions of gains and losses is able to change subjects' degree of loss aversion. In particular, it has been demonstrated that wider ranges of the loss domain leads to lower sensitivity towards losses thus resulting in a reversal of loss aversion. The explanation for this behavior is provided by the Decision by Sampling theory (Stewart et al., 2006): A given loss looks better when most of the other losses being offered are larger. Although beyond the scope of this paper, we hope that future work will consider the impact of differing gain and loss ranges on human choice behavior in sequential search tasks.

4.3.4.3 *Basic tendency for risk seeking preference in MTurk Samples*

We find that our sample of participants are overall slightly more risk seeking (in the gain domain) by being more likely to prefer gambles with lower expected value compared to payoffs with certainty. However, the general tendency of risk aversion is a very common and robust phenomenon: when people make a choice between a risky and a certain reward with identical expected values they tend to prefer the safe option (Kahneman et al., 2013a, 2013b). Moreover, a study investigating risk preferences of MTurk participants has showed that this sample is even slightly more risk–avers compared to a student population (Goodman et al., 2013). Our results do not agree with these studies and we conclude that either the description format of outcomes and probabilities (see Fig. 4.5 D) or the process of spinning a wheel of fortune causes a bias towards choose the gamble above the certain outcome.

4.4 GENERAL DISCUSSION

In an optimal stopping task, values are presented sequentially and only one option can be selected. A consistent finding reveals that individual's behavior relative to the baseline of expected value maximization is inconsistent across search, shifting from risk averse to risk seeking behavior. However, the architecture of the task involves different features of the decision environment, such as stepwise incremental decisions, an unequal frequency of reject decision or inferring the underlying distribution by experience. We found that specific characteristics of the task override stable risk preferences leading to the apparent inconsistent risky behavior across the course of search. Consequently, the lack in consistency of observed risk preferences between sequential decisions and single gambles (Frey et al., 2017; Pedroni et al., 2017) can partly be attributed to peoples' adaptation to the features of the sequential tasks which takes precedent over any general tendency of risk preferences.

Nevertheless, we found small but significant correlations between risk preferences measured in the gamble task and search length in the optimal stopping task (r 0.29, p=0.03), which indicates that higher risk seeking preferences go in hand with more search in the in the optimal stopping task. This result is surprising given that previous studies (Frey et al., 2017; Pedroni et al., 2017) could not find any association in risk preferences between single and dynamic risk taking tasks. A possible explanation for these controversial findings is that the particular implementation of the dynamic task may determine to great extent decision behavior and thus the identifiablity of stable risk preferences. Previous studies have used particular implementations of the balloon analogue risk task (BART Lejuez et al., 2002) and the Columbia Card Task (CCT Figner et al., 2009), involving a trial–to–trial change of the key statistical properties from trial to trial (e.g. in the BART, the random determination of the balloon's explosion point). Accordingly, these tasks involve an additional high degree of uncertainty, which may add increased noise on people's decision strategy and thus on the observed risk preferences. The optimal stopping task used here represents a realistic search problem and entails a stable environment across trials, which lets people establish a consistent decision strategy across contexts (Baumann et al., 2020). Under such conditions, observed behavior allows to determine risk preferences and thus give hope to find a general underlying mechanism (Pedroni et al., 2017) of risk attitude across single and sequential risky decision making tasks.

FUTURE RESEARCH

Optimal stopping tasks involve stepwise incremental decisions and immediate feedback, which correspond to "hot" decision making paradigms. Figner et al. (Figner et al., 2009) have shown that in a "hot" version of a dynamic risk taking decision task, the affective system tends to override the deliberative system in states of heightened emotional arousal. Therefore, affective processes may play an important role during the course of search, in which each time step involves an increase of risk. Indeed, this paper has shown that the removal of the stepwise decisions leads to a change in peoples' decision strategy closer to the optimal solution, indicating more deliberative information processing takes control. Therefore, an important goal of future research would be to identify the interplay between deliberative and affective processes during the search for the optimal alternative.

Despite research reporting no convergence in behavioral measurements between "hot" decision making tasks such as the BART and monetary gambles, there is increased effort to examine the presence and directionality of associations between brain activation in such risk task. Results suggest that the low correlation between risk taking measured in these two task is mirrored in limited functional convergence in neural risk matrix regions. However, these results are difficult to interpret since reliability estimates are based on certain BART implementations (i.e., feedback vs. no feedback, total amount of balloons, distribution over explosion points). The optimal stopping paradigm used here reflects tasks closer to reality but reducing uncertainty by providing stable environments. Moreover, it has been shown that peoples decision strategy is consistent across contexts, such as varying numbers of alternatives or changing variances (Baumann et al., submitted). Furthermore, this paper indicates correlations between search length and risk preferences elicited in single gamble task thus providing evidence for an underlying construct of risk. We thus suggest that optimal stopping task used here may be a promising candidate to examine brain function in response to sequential search behavior and might shed further light on the associations between sequential and single gamble tasks.

4.5 SUPPORTING MATERIAL: THE IMPACT OF SEQUENTIAL SEARCH TASK CHARACTERISTICS ON RISK PREFERENCES

Text 1: Determination of wheels of fortune

The values are drawn from a standard normal distribution with a mean of 160 and a standard deviation of 20. Therefore, the density is

$$f(x) = \frac{1}{20\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-160}{20}\right)^2} \tag{4.7}$$

For the calculation of the slices, we truncated the distribution between 120 and 200. *Position 9*

On the 9*th* position, the decision maker can either accept the certain outcome or spin the wheel, which corresponds to a draw from the distribution. Therefore, the slices of the wheel are proportional to the probabilities to draw a value in the indicated ranges. For example, the probability to receive a value that lies between 190 and 200 is

$$P^{Pos9}_{190-200} = 1 - \int_{120}^{190} f(x)\, dx \tag{4.8}$$

Accordingly, the probability for a value between 180 and 190 is the following:

$$P_{180-190}^{Pos9} = 1 - \int_{120}^{180} f(x)\,dx - P_{190-200}^{Pos9} \tag{4.9}$$

Position 7

At the 7*th* position, one can draw up to 3 times from the distribution. In this case, the sizes of the slides corresponding to the value ranges change as follows:

The probability to receive a value in the range of 190 - 200, under the condition that one can draw up to 3 times, is:

$$P_{190-200}^{Pos7} = 1 - \left(\int_{120}^{190} f(x)\,dx\right)^3 \tag{4.10}$$

The probability to receive a value in the range of 180 - 190, under the condition that one can draw up to 3 times, is:

$$P_{180-190}^{Pos7} = 1 - \left(\int_{120}^{180} f(x)\,dx\right)^3 - P_{190-200}^{Pos7} \tag{4.11}$$

etc.

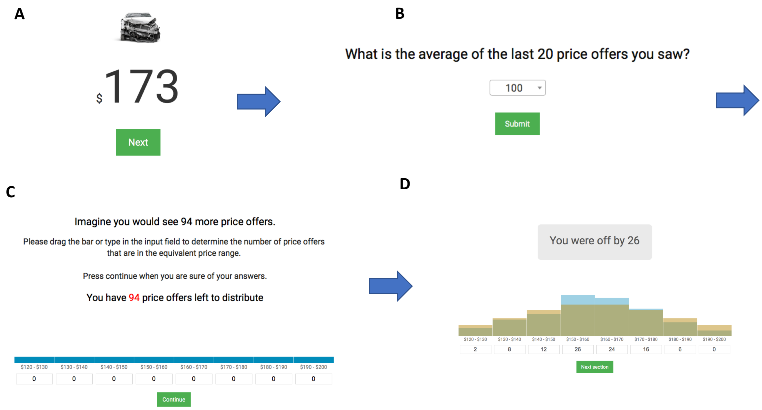

FIGURE 4.8: Learning phase A: Sequential presentation of price offers sampled from predefined distribution 160, 20 . B: After each 20 prices, participants are asked to estimate the average of the prices just seen. C: At the end of the learning phase, participants have to predict how a future sample from the same predefined population might look, where they essentially had to draw a histogram using this interface. D: Feedback was provided by superimposing the correct distribution over their estimate.

BIBLIOGRAPHY

Alós-Ferrer, C., Hügelschäfer, S., & Li, J. (2016). Inertia and decision making. *Frontiers in psychology*, *7*, 169.

Barron, G., & Erev, I. (2003). Small feedback-based decisions and their limited correspondence to description-based decisions. *Journal of behavioral decision making*, *16*(3), 215.

Baumann, C., Schlegelmilch, R., & von Helversen, B. (submitted). Adaptive behavior in optimal sequential search.

Baumann, C., Singmann, H., Gershman, S. J., & von Helversen, B. (2020). A linear threshold model for optimal stopping behavior. *Proceedings of the national academy of sciences*, *117*(23), 12750.

Bavard, S., Lebreton, M., Khamassi, M., Coricelli, G., & Palminteri, S. (2018). Reference-point centering and range-adaptation enhance human reinforcement learning at the cost of irrational preferences. *Nature communications*, *9*(1), 1.

Beach, L. R., & Mitchell, T. R. (1978). A contingency model for the selection of decision strategies. *Academy of management review*, *3*(3), 439.

Bearden, N., Murphy, R., & Rapoport, A. (2005). A multi-attribute extension of the secretary problem: Theory and experiments. *Journal of mathematical psychology*, *49*(5), 410.

Bearden, N., Rapoport, A., & Murphy, R. (2006). Experimental studies of sequential selection and assignment with relative ranks. *Journal of behavioral decision making*, *19*(3), 229.

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50*(1-3), 7.

Bechara, A., Damasio, H., & Damasio, A. R. (1999). Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. *Journal of neuroscience*, *19*(13), 5473.

Benartzi, S., & Thaler, R. H. (1999). Risk aversion or myopia? choices in repeated gambles and retirement investments. *Management science*, *45*(3), 364.

Bernoulli, D. (2011). Exposition of a new theory on the measurement of risk. *The kelly capital growth investment criterion: Theory and practice* (pp. 11–24). World Scientific.

Bhatia, S., He, L., Zhao, W. J., & Analytis, P. P. (2021). Cognitive models of optimal sequential search with recall. *Cognition*, *210*, 104595.

Birnbaum, M. H. (1992). Violations of monotonicity and contextual effects in choice-based certainty equivalents. *Psychological science*, *3*(5), 310.

Bishara, A. J., Pleskac, T. J., Fridberg, D. J., Yechiam, E., Lucas, J., Busemeyer, J. R., Finn, P. R., & Stout, J. C. (2009). Similar processes despite divergent behavior in two commonly used measures of risky decision making. *Journal of behavioral decision making*, *22*(4), 435.

Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., & White, J.-S. S. (2009). Generalized linear mixed models: A practical guide for ecology and evolution. *Trends in ecology & evolution*, *24*(3), 127.

Brehmer, B. (1992). Dynamic decision making: Human control of complex systems. *Acta psychologica*, *81*(3), 211.

Brickman, P. (1972). Optional stopping on ascending and descending series. *Organizational behavior and human performance*, *7*(1), 53.

Brown, G. D., Gardner, J., Oswald, A. J., & Qian, J. (2008). Does wage rank affect employees' well-being? *Industrial relations: a journal of economy and society*, *47*(3), 355.

Busemeyer, J. R., & Stout, J. C. (2002). A contribution of cognitive decision models to clinical assessment: Decomposing performance on the bechara gambling task. *Psychological assessment*, *14*(3), 253.

Camerer, C. (1998). Bounded rationality in individual decision making. *Experimental economics*, *1*(2), 163.

Charness, G., & Levin, D. (2005). When optimal choices feel wrong: A laboratory study of bayesian updating, complexity, and affect. *American economic review*, *95*(4), 1300.

Corbin, R. M., Olson, C. L., & Abbondanza, M. (1975). Context effects in optional stopping decisions. *Organizational behavior and human performance*, *14*(2), 207.

Cox, J. C., & Oaxaca, R. L. (1989). Laboratory experiments with a finite-horizon job-search model. *Journal of risk and uncertainty*, *2*(3), 301.

Dudey, T., & Todd, P. M. (2001). Making good decisions with minimal information: Simultaneous and sequential choice. *Journal of bioeconomics*, *3*(2), 195.

Edwards, W. (1962). Dynamic decision theory and probabilistic information processings. *Human factors*, *4*(2), 59.

Erev, I. (2012). On surprise, change, and the effect of recent outcomes. *Frontiers in psychology*, *3*, 24.

Erev, I., & Haruvy, E. (2016). 10. learning and the economics of small decisions. *The handbook of experimental economics, volume 2* (pp. 638–716). Princeton University Press.

Ferguson, T. (1989). Who solved the secretary problem? *Statistical science*, *4*(3), 282.

Figner, B., Mackinlay, R. J., Wilkening, F., & Weber, E. U. (2009). Affective and deliberative processes in risky choice: Age differences in risk taking in the columbia card task. *Journal of experimental psychology: learning, memory, and cognition*, *35*(3), 709.

Frey, R., Pedroni, A., Mata, R., Rieskamp, J., & Hertwig, R. (2017). Risk preference shares the psychometric structure of major psychological traits. *Science advances*, *3*(10), e1701381.

Genest, W., Stauffer, W. R., & Schultz, W. (2016). Utility functions predict variance and skewness risk preferences in monkeys. *Proceedings of the national academy of sciences*, *113*(30), 8402.

Gigerenzer, G., & Brighton, H. (2009). Homo heuristicus: Why biased minds make better inferences. *Topics in cognitive science*, *1*(1), 107.

Gigerenzer, G., & Selten, R. (2002). *Bounded rationality: The adaptive toolbox*. MIT press.

Gilbert, J. P., & Mosteller, F. (1966). Recognizing the maximum of a sequence. *Journal of the american statistical association*, *61*(313), 35.

Gilovich, T., Vallone, R., & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive psychology*, *17*(3), 295.

Goldstein, D. G., McAfee, R. P., Suri, S., & Wright, J. R. (2020). Learning when to stop searching. *Management science*, *66*(3), 1375.

Goldstein, D. G., & Rothschild, D. (2014). Lay understanding of probability distributions. *Judgment and decision making*, *9*(1).

Gonzalez, C., Fakhari, P., & Busemeyer, J. (2017). Dynamic decision making: Learning processes and new research directions. *Human factors*, *59*(5), 713.

Goodman, J. K., Cryder, C. E., & Cheema, A. (2013). Data collection in a flat world: The strengths and weaknesses of mechanical turk samples. *Journal of behavioral decision making*, *26*(3), 213.

Guan, M., & Lee, M. (2018). The effect of goals and environments on human performance in optimal stopping problems. *Decision*, *5*(4), 339.

Guan, M., Lee, M., & Silva, A. (2014). Threshold models of human decision making on optimal stopping problems in different environments. In P. Bello, M. M. M. Guarini, & B. Scassellati (Eds.), *Proceedings of the annual meeting of the cognitive science society* (pp. 553–558).

Guan, M., Lee, M., & Vandekerckhove, J. (2015). Threshold Models of Human Decision Making on Optimal Stopping Problems in Different Environments (D. Noelle & R. Dale, Eds.). *Proceedings of the 37th annual meeting of the cognitive science society*, *37*(37), 824.

Guan, M., Stokes, R., Vandekerckhove, J., & Lee, M. (2020). A cognitive modeling analysis of risk in sequential choice tasks. *Judgment and decision making*, *15*, 823.

Hayden, B. Y., Heilbronner, S. R., Nair, A. C., & Platt, M. L. (2008). Cognitive influences on risk-seeking by rhesus macaques. *Judgment and decision making*, *3*(5), 389.

Heilbronner, S., & Hayden, B. (2013). Contextual factors explain risk-seeking preferences in rhesus monkeys. *Frontiers in neuroscience*, *7*, 7.

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological science*, *15*(8), 534.

Hertwig, R., & Erev, I. (2009). The description–experience gap in risky choice. *Trends in cognitive sciences*, *13*(12), 517.

Hey, J. D. (1982). Search for rules for search. *Journal of economic behavior & organization*, *3*(1), 65.

Hills, T. T., & Pachur, T. (2012). Dynamic search and working memory in social recall. *Journal of experimental psychology: learning, memory, and cognition*, *38*(1), 218.

Hills, T. T., Todd, P. M., & Goldstone, R. L. (2008). Search in external and internal spaces: Evidence for generalized cognitive search processes. *Psychological science*, *19*(8), 802.

Hills, T. T., Todd, P. M., Lazer, D., Redish, A. D., Couzin, I. D., Group, C. S. R., et al. (2015). Exploration versus exploitation in space, mind, and society. *Trends in cognitive sciences*, *19*(1), 46.

Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effects. *American economic review*, *92*(5), 1644.

Huys, Q. J., Lally, N., Faulkner, P., Eshel, N., Seifritz, E., Gershman, S. J., Dayan, P., & Roiser, J. P. (2015). Interplay of approximate planning strategies. *Proceedings of the national academy of sciences*, *112*(10), 3098.

Imas, A. (2016). The realization effect: Risk-taking after realized versus paper losses. *American economic review*, *106*(8), 2086.

Jessup, R. K., Bishara, A. J., & Busemeyer, J. R. (2008). Feedback produces divergence from prospect theory in descriptive choice. *Psychological science*, *19*(10), 1015.

Kacelnik, A., & Bateson, M. (1996). Risky theories—the effects of variance on foraging decisions. *American zoologist*, *36*(4), 402.

Kahan, J. P., Rapoport, A., & Jones, L. V. (1967). Decision making in a sequential search task. *Perception and psychophysics*, *2*(8), 374.

Kahneman, D., & Tversky, A. (1979). On the interpretation of intuitive probability: A reply to jonathan cohen. *Cognition*, *7*(4), 409.

Kahneman, D., & Tversky, A. (2013a). Choices, values, and frames. *Handbook of the fundamentals of financial decision making: Part i* (pp. 269–278). World Scientific.

Kahneman, D., & Tversky, A. (2013b). Prospect theory: An analysis of decision under risk. *Handbook of the fundamentals of financial decision making: Part i* (pp. 99–127). World Scientific.

Kalish, M. L., Griffiths, T. L., & Lewandowsky, S. (2007). Iterated learning: Intergenerational knowledge transmission reveals inductive biases. *Psychonomic bulletin & review*, *14*(2), 288.

Keren, G., & Lewis, C. (1994). The two fallacies of gamblers: Type i and type ii. *Organizational behavior and human decision processes*, *60*(1), 75.

Klein, T. A., Ullsperger, M., & Jocham, G. (2017). Learning relative values in the striatum induces violations of normative decision making. *Nature communications*, *8*, 16033.

Klos, A., Weber, E. U., & Weber, M. (2005). Investment decisions and time horizon: Risk perception and risk behavior in repeated gambles. *Management science*, *51*(12), 1777.

Kogut, C. A. (1990). Consumer search behavior and sunk costs. *Journal of economic and behavioral organization*, *14*(3), 381.

Lee, M. (2006). A hierarchical bayesian model of human decision-making on an optimal stopping problem. *Cognitive science*, *30*(3), 1.

Lee, M., O'Connor, T., & Welsh, M. (2004). Threshold models of human decision making on optimal stopping problems in different environments. In K. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the 26th annual conference of the cognitive science society* (pp. 819–824).

Lee, M., & Wagenmakers, E.-J. (2014). *Bayesian cognitive modeling: A practical course*. Cambridge university press.

Lejarraga, T., & Gonzalez, C. (2011). Effects of feedback and complexity on repeated decisions from description. *Organizational behavior and human decision processes*, *116*(2), 286.

Lejuez, C. W., Read, J. P., Kahler, C. W., Richards, J. B., Ramsey, S. E., Stuart, G. L., Strong, D. R., & Brown, R. A. (2002). Evaluation of a behavioral measure of risk taking: The balloon analogue risk task. *Journal of experimental psychology: applied*, *8*(2), 75.

Little, D. R., & Shiffrin, R. (2009). Simplicity bias in the estimation of causal functions. *Proceedings of the annual meeting of the cognitive science society*, *31*(31).

Louie, K., & Glimcher, P. W. (2012). Efficient coding and the neural representation of value. *Annals of the new york academy of sciences*, *1251*(1), 13.

Lucas, C. G., Griffiths, T. L., Williams, J. J., & Kalish, M. L. (2015). A rational model of function learning. *Psychonomic bulletin & review*, *22*(5), 1193.

Ludvig, E. A., Madan, C. R., McMillan, N., Xu, Y., & Spetch, M. L. (2018). Living near the edge: How extreme outcomes and their neighbors drive risky choice. *Journal of experimental psychology: general*, *147*(12), 1905.

Madan, C. R., Ludvig, E. A., & Spetch, M. L. (2019). Comparative inspiration: From puzzles with pigeons to novel discoveries with humans in risky choice. *Behavioural processes*, *160*, 10.

Markowitz, H. (1959). *Portfolio selection: Efficient diversification of investments* (Vol. 16). John Wiley New York.

Mata, R., & von Helversen, B. (2015). Search and the aging mind: The promise and limits of the cognitive control hypothesis of age differences in search. *Topics in cognitive science*, *7*(3), 416.

Mata, R., Wilke, A., & Czienskowski, U. (2013). Foraging across the life span: Is there a reduction in exploration with aging? *Frontiers in neuroscience*, *7*, 53.

McCoy, A. N., & Platt, M. L. (2005). Expectations and outcomes: Decision-making in the primate brain. *Journal of comparative physiology a*, *191*(3), 201.

McGlothlin, W. H. (1956). Stability of choices among uncertain alternatives. *The american journal of psychology*, *69*(4), 604.

Meyers, E. A., & Koehler, D. J. (2020). Individual differences in exploring versus exploiting and links to delay discounting. *Journal of behavioral decision making*.

Miler, K. et al. (2009). Doomed to repeat the successes of the past: History is best forgotten for repeated choices with nonstationary payoffs. *Memory & cognition*, *37*(7), 985.

Moorthy, S., Ratchford, B. T., & Talukdar, D. (1997). Consumer information search revisited: Theory and empirical analysis. *Journal of consumer research*, *23*(4), 263.

Morey, R., Rouder, J., Jamil, T., Urbanek, S., Forner, K., & Ly, A. (2018). *Bayesfactor: Computation of bayes factors for common designs* [R package version 0.9.12-4.2].

Newell, A., Simon, H. A. et al. (1972). *Human problem solving* (Vol. 104). Prentice-Hall Englewood Cliffs, NJ.

Nieuwenhuis, S., Heslenfeld, D. J., von Geusau, N. J. A., Mars, R. B., Holroyd, C. B., & Yeung, N. (2005). Activity in human reward-sensitive brain areas is strongly context dependent. *Neuroimage*, *25*(4), 1302.

Ofek, E., Yildiz, M., & Haruvy, E. (2007). The impact of prior decisions on subsequent valuations in a costly contemplation model. *Management science*, *53*(8), 1217.

Padoa-Schioppa, C. (2009). Range-adapting representation of economic value in the orbitofrontal cortex. *Journal of neuroscience*, *29*(44), 14004.

Payne, J. W. (1982). Contingent decision behavior. *Psychological bulletin*, *92*(2), 382.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of experimental psychology: learning, memory, and cognition*, *14*(3), 534.

Payne, J. W., Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge university press.

Paz, L., Insabato, A., Zylberberg, A., Deco, G., & Sigman, M. (2016). Confidence through consensus: A neural mechanism for uncertainty monitoring. *Scientific reports*, *6*(1), 1.

Pedroni, A., Frey, R., Bruhin, A., Dutilh, G., Hertwig, R., & Rieskamp, J. (2017). The risk elicitation puzzle. *Nature human behaviour*, *1*(11), 803.

Pirolli, P. (2007). *Information foraging theory: Adaptive interaction with information*. Oxford University Press.

Pleskac, T. J. (2008). Decision making and learning while taking sequential risks. *Journal of experimental psychology: learning, memory, and cognition*, *34*(1), 167.

Plummer, M. et al. (2003). Jags: A program for analysis of bayesian graphical models using gibbs sampling. *Proceedings of the 3rd international workshop on distributed statistical computing*, *124*(125.10).

Rabin, M., & Weizsäcker, G. (2009). Narrow bracketing and dominated choices. *American economic review*, *99*(4), 1508.

Rangel, A., & Clithero, J. A. (2012). Value normalization in decision making: Theory and evidence. *Current opinion in neurobiology*, *22*(6), 970.

Rapoport, A., & Tversky, A. (1966). Cost and accessibility of offers as determinants of optional stopping. *Psychonomic science*, *4*(1), 145.

Rapoport, A., & Tversky, A. (1970). Choice behavior in an optional stopping task. *Organ behav hum decis process*, *5*(2), 105.

Read, D., Loewenstein, G., Rabin, M., Keren, G., & Laibson, D. (1999). Choice bracketing. *Elicitation of preferences* (pp. 171–202). Springer.

Rigoli, F., Friston, K. J., & Dolan, R. J. (2016). Neural processes mediating contextual influences on human choice behaviour. *Nature communications*, *7*(1), 1.

Russo, J. E., & Dosher, B. A. (1983). Strategies for multiattribute binary choice. *Journal of experimental psychology: learning, memory, and cognition*, *9*(4), 676.

Rydzewska, K., von Helversen, B., Kossowska, M., Magnuski, M., & Sedek, G. (2018). Age-related within-task adaptations in sequential decision making: Considering cognitive and motivational factors. *Psychology and aging*, *33*(2), 297.

Sang, K., Todd, P. M., Goldstone, R. L., & Hills, T. T. (2020). Simple threshold rules solve explore/exploit trade-offs in a resource accumulation search task. *Cognitive science*, *44*(2), e12817.

Schotter, A., & Braunstein, Y. M. (1981). Economic search: An experimental study. *Economic inquiry*, *19*(1), 1.

Schunk, D. (2009). Behavioral heterogeneity in dynamic search situations: Theory and experimental evidence. *Journal of economic dynamics and control*, *33*(9), 1719.

Schunk, D., & Winter, J. (2009). The relationship between risk attitudes and heuristics in search tasks: A laboratory experiment. *Journal of economic behavior & organization*, *71*(2), 347.

Seale, D. A., & Rapoport, A. (1997). Sequential Decision Making with Relative Ranks : An Experimental Investigation of the “Secretary Problem”. *Organizational behavior and human decision processes*, *69*(3), 221.

Seale, D. A., & Rapoport, A. (2000). Optimal stopping behavior with relative ranks: The secretary problem with unknown population size. *Journal of behavioral decision making*, *13*(4), 391.

Shapira, Z., & Venezia, I. (1981). Optional stopping on nonstationary series. *Organizational behavior and human performance*, *27*(1), 32.

Simon, H. A. (1955). A behavioral model of rational choice. *The quarterly journal of economics*, *69*(1), 99.

Simon, H. A. (1957). *Models of man; social and rational.* Wiley.

Simon, H. A. (1990). Bounded rationality. In J. Eatwell, M. Milgate, & P. Newman (Eds.), *Utility and probability* (pp. 15–18). Palgrave Macmillan UK.

Simon, H. A. (1997). *Models of bounded rationality: Empirically grounded economic reason* (Vol. 3). MIT press.

Singmann, H., & Kellen, D. (2019). An introduction to mixed models for experimental psychology. *New methods in cognitive psychology*, *28*, 4.

Song, M., Bnaya, Z., & Ma, W. J. (2019). Sources of suboptimality in a minimalistic explore–exploit task. *Nature human behavior*, 1.

Sonnemans, J. (1998). Strategies of search. *Journal of economic behavior & organization*, *35*(3), 309.

Sonnemans, J. (2000). Decisions and strategies in a sequential search experiment. *Journal of economic psychology*, *21*(1), 91.

Sonsino, D., Benzion, U., & Mador, G. (2002). The complexity effects on choice with uncertainty–experimental evidence. *The economic journal*, *112*(482), 936.

Speekenbrink, M., & Konstantinidis, E. (2015). Uncertainty and exploration in a restless bandit problem. *Topics in cognitive science*, *7*(2), 351.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the royal statistical society: series b (statistical methodology)*, *64*(4), 583.

Stango, V., & Yinman, J. (2009). Exponential growth bias and household finance. *The journal of finance*, *64*(6), 2807.

Stewart, N., Chater, N., & Brown, G. D. (2006). Decision by sampling. *Cognitive psychology*, *53*(1), 1.

Stewart, N., Chater, N., Stott, H. P., & Reimers, S. (2003). Prospect relativity: How choice options influence decision under risk. *Journal of experimental psychology: general*, *132*(1), 23.

Steyvers, M., Lee, M., & Wagenmakers, E.-J. (2009). A bayesian analysis of human decision-making on bandit problems. *Journal of mathematical psychology*, *53*(3), 168.

Stigler, G. J. (1961). The economics of information. *Journal of political economy*, *69*(3), 213.

Stout, J. C., Busemeyer, J. R., Lin, A., Grant, S. J., & Bonson, K. R. (2004). Cognitive modeling analysis of decision-making processes in cocaine abusers. *Psychonomic bulletin & review*, *11*(4), 742.

Thaler, R. H., & Johnson, E. J. (1990). Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice. *Management science*, *36*(6), 643.

Todd, P. M. (2001). Fast and frugal heuristics for environmentally bounded minds. *Bounded rationality: the adaptive toolbox*, 51.

Todd, P. M., & Gigerenzer, G. E. (2012). *Ecological rationality: Intelligence in the world.* Oxford University Press.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*(4481), 453.

Tversky, A., & Simonson, I. (1993). Context-dependent preferences. *Management science*, *39*(10), 1179.

von Helversen, B., & Mata, R. (2012). Losing a dime with a satisfied mind: Positive affect predicts less search in sequential decision making. *Psychology and aging*, *27*(4), 825.

von Helversen, B., Mata, R., Samanez-Larkin, G. R., & Wilke, A. (2018). Foraging, exploration, or search? on the (lack of) convergent validity between three behavioral paradigms. *Evolutionary behavioral sciences*, *12*(3), 152.

von Helversen, B., Wilke, A., Johnson, T., Schmid, G., & Klapp, B. (2011). Performance benefits of depression: Sequential decision making in a healthy sample and a clinically depressed sample. *Journal of abnormal psychology*, *120*(4), 962.

Wagenaar, W. A., & Sagaria, S. D. (1975). Misperception of exponential growth. *Percept psychophys*, *18*(6), 416.

Walasek, L., & Stewart, N. (2015). How to make loss aversion disappear and reverse: Tests of the decision by sampling origin of loss aversion. *Journal of experimental psychology: general*, *144*(1), 7.

Wallsten, T. S., Pleskac, T. J., & Lejuez, C. W. (2005). Modeling behavior in a clinically diagnostic sequential risk-taking task. *Psychological review*, *112*(4), 862.

Weber, E., Shafir, S., & Blais, A.-R. (2004). Predicting risk sensitivity in humans and lower animals: Risk as variance or coefficient of variation. *Psychological review*, *111*, 430.

Weber, E. U., & Johnson, E. J. (2009). Mindful judgment and decision making. *Annual review of psychology*, *60*, 53.

Weber, M., & Zuchel, H. (2005). How do prior outcomes affect risk attitude? comparing escalation of commitment and the house-money effect. *Decision analysis*, *2*(1), 30.

Wilson, R. C., Geana, A., White, J. M., Ludvig, E. A., & Cohen, J. D. (2014). Humans use directed and random exploration to solve the explore–exploit dilemma. *Journal of experimental psychology: general*, *143*(6), 2074.

Wulff, D. U., Mergenthaler-Canseco, M., & Hertwig, R. (2018). A meta-analytic review of two modes of learning and the description-experience gap. *Psychological bulletin*, *144*(2), 140.

Zwick, R., Rapoport, A., Lo, A. K. C., & Muthukrishnan, a. V. (2003). Consumer Sequential Search: Not Enough or Too Much? *Marketing science*, *22*(4), 503.